1982

Estimation of random effects in the balanced one-way classification

Julio León Peixoto
Iowa State University

8307777

Peixoto, Julio Leon

ESTIMATION OF RANDOM EFFECTS IN THE BALANCED ONE-WAY CLASSIFICATION

Iowa State University PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Estimation of random effects in the balanced one-way classification

by

Julio León Peixoto

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

I. INTRODUCTION

A. Introduction to the One-way Random Model

Searle (1971, p. 377), discussed a laboratory experiment to study the maternal ability of mice. This experiment used litter weights of ten-day-old litters as a measure of maternal ability. Six litters from each of four dams, all of one breed, constituted the experimental units. A suitable model for analyzing the data is the one-way classification model

$$y_{ij} = \mu + a_i + e_{ij} \qquad (i = 1,2,3,4;\ j = 1,2,3,4,5,6),$$

where y_{ij} is the weight of the jth litter from the ith dam, μ represents the overall mean, a_i is the effect due to the ith dam, and e_{ij} is an error term.

Maternal ability is surely a variable that is subject to biological variation from animal to animal. The prime concern of the experiment is unlikely to center specifically on the four female mice used in the experiment. They can be regarded as a sample of size four from a very large population of female mice.

The model for this experiment is a particular case of the <u>balanced one-way random model</u>:

$$y_{ij} = \mu + a_i + e_{ij} \qquad (i = 1,\ldots,I;\ \ j = 1,\ldots,J),$$

where y_{ij} is the response of the jth unit in the ith group, μ

is the overall mean, a_i is a random effect associated with the ith group, and e_{ij} is a random error or residual effect. We assume that $a_1,\ldots,a_I$ constitutes a random sample from a normal distribution with zero mean and variance σ_a^2. Similarly, $e_{11},e_{12},\ldots,e_{IJ}$ are assumed to be a random sample from a normal distribution with mean zero and variance σ_e^2. Furthermore, sampling of the a_i's is assumed to be independent of that of the e_{ij}'s, implying in particular that the covariance between each a_i and each e_{ij} is zero.

The balanced one-way random model can be summarized as follows:

$$y_{ij} = \mu + a_i + e_{ij} \quad (i = 1,\ldots,I; \;\; j = 1,\ldots,J), \qquad \text{(I.A.1)}$$

where $a_1,\ldots,a_I$ are identically distributed as $N(0,\sigma_a^2)$, $e_{11},e_{12},\ldots,e_{IJ}$ are identically distributed as $N(0,\sigma_e^2)$, and $a_1,\ldots,a_I,e_{11},e_{12},\ldots,e_{IJ}$ are statistically independent.

Let α_i be the realized, but unobservable, value of the random effect a_i. In many applications, there is interest in estimating the α_i's. The estimation of α_i is sometimes called "prediction of a_i." For the mice example, one objective of the experiment may have been to evaluate the four females as candidates for some subsequent experiment.

The subject of this dissertation is statistical inference for μ, for the α_i's, and for linear combinations of these parameters. In most past work on this problem, it has been assumed that the variance ratio σ_a^2/σ_e^2 is known. However, this assumption is generally not realistic. We shall describe various estimators of the parameters of interest and study certain properties of these estimators, taking the variance ratio to be unknown. We consider both unconditional and conditional properties. By unconditional properties we mean properties based on model (I.A.1). By conditional properties, we mean properties based on model (I.A.1), but conditional on $a_i = \alpha_i$ $(i = 1,\ldots,I)$. Throughout this dissertation we use the terms bias, total bias, mean squared error, and total mean squared error as equivalent to the terms unconditional bias, total unconditional bias, unconditional mean squared error, and total unconditional mean squared error, respectively. In essence, model (I.A.1) conditional on $a_i = \alpha_i$ $(i = 1,\ldots,I)$ is the balanced one-way fixed model, which can be written as

$$y_{ij} = \mu + \alpha_i + e_{ij} \quad (i = 1,\ldots,I;\ \ j = 1,\ldots,J), \qquad \text{(I.A.2)}$$

where $e_{11}, e_{12}, \ldots e_{IJ}$ are distributed independently as $N(0,\sigma_e^2)$ and where $\mu, \alpha_1, \ldots, \alpha_I$ are fixed and unknown. Some estimators of linear combinations of $\mu, \alpha_1, \ldots, \alpha_I$ derived under the random model (I.A.1) will be seen to have good conditional properties, as well as good unconditional properties. In fact, the random model (I.A.1) can be regarded as a Bayesian formulation of the fixed model (I.A.2), so that estimators of linear combinations of $\mu, \alpha_1, \ldots, \alpha_I$ under (I.A.1) can be interpreted as Bayesian estimators of the same parameters under (I.A.2).

We now introduce some notation (in the context of the balanced one-way random model) that will be used in subsequent chapters.

Let

$$\mu_i = \mu + \alpha_i \tag{I.A.3}$$

represent the conditional mean of the ith group, and let

$$m_i = \mu + a_i \quad (i = 1, \ldots, I). \tag{I.A.4}$$

Define the vectors

$$\underline{a} = (a_1, \ldots, a_I)', \ \underline{\alpha} = (\alpha_1, \ldots, \alpha_I)', \tag{I.A.5}$$

$$\underline{\mu} = (\mu_1, \ldots, \mu_I)', \text{ and } \underline{m} = (m_1, \ldots, m_I)'. \tag{I.A.6}$$

The among-groups and within-groups expected mean squares are denoted by

$$\gamma_a = \sigma_e^2 + J\sigma_a^2 \quad \text{and} \quad \gamma_e = \sigma_e^2, \tag{I.A.7}$$

respectively. The ratio of expected mean squares is

$$\rho = \frac{\gamma_e}{\gamma_a} = \frac{\sigma_e^2 + J\sigma_a^2}{\sigma_e^2}. \tag{I.A.8}$$

Note that ρ is known if and only if the variance ratio

σ_a^2/σ_e^2 is known.

Let

$$\mu^* = \mu + \bar{\alpha}_. \text{ and } \alpha_i^* = \alpha_i - \bar{\alpha}_. \qquad (i = 1,\ldots,I), \tag{I.A.9}$$

where

$$\bar{\alpha}_. = (1/I) \sum_{h=1}^{I} \alpha_h. \tag{I.A.10}$$

The among-groups and within-groups sums of squares are

$$SS_a = J\sum_i (\bar{y}_{i.}-\bar{y}_{..})^2, \text{ and } SS_e = \sum_{ij}\sum (y_{ij}-\bar{y}_{..})^2, \tag{I.A.11}$$

respectively. The corresponding mean squares are given by

$$MS_a = SS_a/(I-1), \text{ and } MS_e = SS_e/(IJ-I). \tag{I.A.12}$$

Define

$$s_\alpha^2 = (\sum_i \alpha_i^{*2})/(I-1), \tag{I.A.13}$$

and

$$\lambda = (1/2)(I-1)J(s_\alpha^2/\sigma_e^2). \tag{I.A.14}$$

Let

$$\tau = \ell_0\mu + \sum_{i=1}^{I} \ell_i\alpha_i \tag{I.A.15}$$

represent an arbitrary linear combination of the parameters $\mu,\alpha_1,\ldots,\alpha_I$, and let

$$t = \ell_0\mu + \sum_{i=1}^{I} \ell_i a_i. \tag{I.A.16}$$

The symbols $\hat{\alpha}_i$, $\hat{\mu}_i$, $\hat{\underline{\mu}}$, $\hat{\mu}$, and $\hat{\tau}$ will be used to denote arbitrary estimators of α_i, μ_i, $\underline{\mu}$, μ, and τ, respectively, unless otherwise stated. The symbol $\underline{y}$ denotes the vector of responses $(y_{11}, y_{12}, \ldots, y_{IJ})'$.

The one-way random model (I.A.1) is a particular case of the general mixed linear model. It is convenient to present some basic definitions and properties in the context of this general model.

B. The General Mixed Model: Basic Definitions and Results

The <u>general mixed model</u> is:

$$\underline{y} = \underline{X\beta} + \underline{Za} + \underline{e}, \qquad \text{(I.B.1)}$$

where $\underline{y}$ is an nx1 vector of random variables whose observed values comprise the data points, $\underline{X}$ and $\underline{Z}$ are matrices of known "regressors" with dimensions nxp and nxq, respectively, $\underline{\beta}$ is a px1 vector of fixed unknown parameters, and $\underline{a}$ and $\underline{e}$ are random vectors of dimensions qx1 and nx1, respectively. It is assumed that $\underline{a}$ is distributed as $N(\underline{0}, \sigma_e^2\underline{D})$, $\underline{e}$ is distributed as $N(\underline{0}, \sigma_e^2\underline{I})$, and that $\underline{a}$ and $\underline{e}$ are statistically independent. The symbol $\underline{I}$ represents the identity matrix, $\underline{D}$ is a nonnegative definite matrix, not necessarily known, and σ_e^2 is an unknown positive parameter.

Let $\underline{\alpha}$ represent the realized value of the random vector

$\underline{a}$. Our objective is to estimate linear combinations of the elements of $\underline{\beta}$ and $\underline{\alpha}$. Let $\tau = \underline{\lambda}_1'\underline{\beta} + \underline{\lambda}_2'\underline{\alpha}$ represent an arbitrary linear combination of these elements, and let $t = \underline{\lambda}_1'\underline{\beta} + \underline{\lambda}_2'\underline{a}$.

The term conditional will subsequently be used to mean "conditional on $\underline{a} = \underline{\alpha}$", unless otherwise stated. Note that the general mixed linear model conditional on $\underline{a} = \underline{\alpha}$ is a <u>fixed linear model</u>.

<u>Definition I.B.1</u>: An estimator $\hat{\tau}(\underline{y})$ is said to be <u>linear</u> if $\hat{\tau}(\underline{y}) = c+\underline{d}'\underline{y}$ for some constant c and some constant vector $\underline{d}$.

<u>Definition I.B.2</u>: Under the general mixed linear model, an estimator $\hat{\tau}(\underline{y})$ of τ is said to be <u>unbiased</u> (or <u>unconditionally unbiased</u>) if $E[\hat{\tau}(\underline{y})] = E(t) = \underline{\lambda}_1'\underline{\beta}$.

<u>Definition I.B.3</u>: Under the general mixed linear model, an estimator $\hat{\tau}(\underline{y})$ of τ is said to be <u>conditionally unbiased</u> if $E[\hat{\tau}(\underline{y})|\underline{a} = \underline{\alpha}] = \tau$.

<u>Definition I.B.4</u>: Under the general mixed linear model, the parametric function τ is said to be estimable if there exists a linear estimator that estimates it unconditionally unbiasedly, or, equivalently, if $\underline{\lambda}_1' = \underline{d}'\underline{X}$ for some constant vector $\underline{d}$.

Definition I.B.5: Under the general mixed linear model, the parametric function τ is said to be conditionally estimable if there exists a linear estimator that estimates it conditionally unbiasedly, or, equivalently, if $\underline{\lambda}_1' = \underline{d}'X$ and $\underline{\lambda}_2' = \underline{d}'Z$ for some constant vector $\underline{d}$.

Note that an estimator that is conditionally unbiased is also unbiased, while the converse is not necessarily true. Similarly, a conditionally estimable parametric function is also estimable, but an estimable function is not necessarily conditionally estimable.

Definition I.B.6: Under the general mixed linear model, the mean squared error (MSE), or unconditional mean squared error, of an estimator $\tau(\underline{y})$ of τ is defined to be $\mathrm{MSE}[\hat{\tau}(\underline{y}),\tau] = E[\hat{\tau}(\underline{y})-t]^2$.

Definition I.B.7: Under the general mixed linear model, the conditional mean squared error (CMSE) of an estimator $\hat{\tau}(\underline{y})$ of τ is defined to be $\mathrm{CMSE}[\hat{\tau}(\underline{y}), \tau] = E\{[\hat{\tau}(\underline{y})-\tau]^2 | \underline{a} = \underline{\alpha}\}$.

It is instructive to think of the CMSE in terms of the role that it plays in best linear unbiased estimation under a fixed linear model. The linear conditionally unbiased estimator that minimizes the CMSE under the general mixed linear model is given by the well-known Gauss-Markov theorem. There is an extended version of this theorem that gives the linear (unconditionally) unbiased estimator that minimizes the (unconditional) MSE under the general mixed linear model. Prior to reviewing these results, it is convenient to review the definitions of normal equations and of equations known as mixed-model equations.

<u>Definition I.B.8</u>: The linear equations

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} \end{bmatrix} \begin{bmatrix} \tilde{\underline{\beta}} \\ \tilde{\underline{\alpha}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix} \qquad \text{(I.B.2)}$$

in $\tilde{\underline{\beta}}$ and $\tilde{\underline{\alpha}}$ are known as the <u>normal equations</u>.

<u>Definition I.B.9</u>: If $\underline{D}$ is known and positive definite, the linear equations

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{D}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\underline{\beta}} \\ \hat{\underline{\alpha}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix} \qquad \text{(I.B.3)}$$

in $\hat{\underline{\beta}}$ and $\hat{\underline{\alpha}}$ are known as (Henderson's) <u>mixed-model equations</u>.

Harville (1976) provides various extended versions of the mixed-model equations for the case where $\underline{D}$ is known and nonnegative definite.

Define

$$\underline{V} = \underline{I} + \underline{ZDZ}' \qquad \text{(I.B.4)}$$

so that (under the general mixed linear model)

$$\text{Var}(\underline{y}) = \sigma_e^2\underline{V}.$$

As discussed in the following theorem, the Equations (I.B.3) can be related to equations introduced by Aitken (1934), known as the Aitken equations.

<u>Theorem I.B.1</u>: The vectors $\hat{\hat{\underline{\beta}}}$ and $\hat{\underline{\alpha}}$ form a solution to Henderson's mixed-model equations if and only if

$$(\underline{X}'\underline{V}^{-1}\underline{X})\hat{\hat{\underline{\beta}}} = \underline{X}'\underline{V}^{-1}\underline{y} \tag{I.B.5}$$

and

$$\hat{\underline{\alpha}} = \underline{DZ}'\underline{V}^{-1}(\underline{y}-\underline{X}\hat{\hat{\underline{\beta}}}). \tag{I.B.6}$$

A proof of this theorem was given, e.g., by Harville (1976). Equations (I.B.5) are the <u>Aitken equations</u>.

Let $\tilde{\tau} = \underline{\lambda}_1'\tilde{\underline{\beta}} + \underline{\lambda}_2'\tilde{\underline{\alpha}}$ and $\hat{\hat{\tau}} = \underline{\lambda}_1'\hat{\hat{\underline{\lambda}}} + \lambda_2'\hat{\hat{\underline{\alpha}}}$.

The following theorem is a restatement (in the context of the general mixed linear model) of the Gauss-Markov theorem.

<u>Theorem I.B.2 (Gauss-Markov)</u>: Take the model to be the general mixed linear model, and let $\tilde{\underline{\beta}}$ and $\tilde{\underline{\alpha}}$ represent any solutions to the normal equations. If τ is conditionally estimable, then $\tilde{\tau}$ is a linear conditionally unbiased estimator of τ and has uniformly smaller CMSE than any other linear conditionally unbiased estimator of τ.

We refer to $\tilde{\tau}$ as the best <u>linear conditionally unbiased estimator</u> (BLCUE) of τ. A proof of the Gauss-Markov theorem is given, e.g., by Graybill (1976, p. 219). This theorem is equivalent to a special case of the following theorem.

Theorem I.B.3 (Extended Gauss-Markov): Take the model to be the general mixed linear model, and let $\hat{\hat{\underline{\beta}}}$ and $\hat{\hat{\underline{\alpha}}}$ represent any solution to the mixed-model equations (I.B.3). If τ is estimable, and $\underline{D}$ is known and positive definite, then $\hat{\hat{\tau}}$ is a linear (unconditionally) unbiased estimator of τ and has uniformly smaller (unconditional) MSE than any other linear unbiased estimator of τ.

We refer to $\hat{\hat{\tau}}$ as the best linear unbiased estimator (BLUE) of τ. A proof of the Extended Gauss-Markov theorem is given, e.g., by Harville (1976), who also provides extended versions for the case where $\underline{D}$ is known and nonnegative definite.

Let $\underline{A}^-$ denote an arbitrary generalized inverse of an arbitrary matrix $\underline{A}$. The following two theorems give the MSE of the BLUE and the CMSE of the BLCUE of a linear parametric function τ.

Theorem I.B.4: Take the model to be the general mixed linear model. If τ is (unconditionally) estimable, and $\underline{D}$ is known and positive definite, then

$$\mathrm{MSE}(\hat{\hat{\tau}}, \tau) = \sigma_e^2[\underline{\lambda}_1', \underline{\lambda}_2'] \begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z}+\underline{D}^{-1} \end{bmatrix}^- \begin{bmatrix} \underline{\lambda}_1 \\ \underline{\lambda}_2 \end{bmatrix} . \qquad \text{(I.B.7)}$$

A proof of Theorem I.B.4 is given, e.g., by Harville (1976), who also provides extended versions for the case where $\underset{\sim}{D}$ is known and nonnegative definite.

Theorem I.B.5: Take the model to be the general mixed linear model. If τ is conditionally estimable, then

$$\text{CMSE}(\tilde{\tau},\tau) = \text{Var}(\tilde{\tau}\,|\,\underset{\sim}{a}=\underset{\sim}{\alpha})$$

$$= \sigma_e^2[\underset{\sim}{\lambda}_1', \underset{\sim}{\lambda}_2'] \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{X} & \underset{\sim}{Z}'\underset{\sim}{Z} \end{bmatrix}^{-} \begin{bmatrix} \underset{\sim}{\lambda}_1 \\ \underset{\sim}{\lambda}_2 \end{bmatrix} . \qquad \text{(I.B.8)}$$

A proof of Theorem I.B.5 is given, e.g., by Searle (1971, Chapter 5). This theorem is equivalent to a special case of Theorem I.B.4.

The concepts of unbiasedness and mean squared error are useful for evaluating the performance of individual estimators. We now give some criteria that can be used to evaluate ensemble properties of a vector of estimators.

Let

$$\underset{\sim}{\mu} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{\alpha} \quad \text{and} \quad \underset{\sim}{m} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{a}. \qquad \text{(I.B.9)}$$

Let $\hat{\underset{\sim}{\mu}} = \hat{\underset{\sim}{\mu}}(\underset{\sim}{y})$ represent an arbitrary estimator of $\underset{\sim}{\mu}$. The ith components of $\hat{\underset{\sim}{\mu}}$, $\underset{\sim}{\mu}$, and $\underset{\sim}{m}$ will be denoted by $\hat{\mu}_i, \mu_i$, and m_i, respectively $(i = 1,\ldots,n)$.

Definition I.B.10: Under the general mixed linear model, the total bias (TB) of $\underline{\hat{\mu}}$ as an estimator of $\underline{\mu}$ is defined as

$$TB(\underline{\hat{\mu}},\underline{\mu}) = \sum_{i=1}^{n} E(\hat{\mu}_i - m_i). \tag{I.B.10}$$

Note that the total bias of $\underline{\hat{\mu}}$ is the sum of the individual biases of its components. An analogous definition for conditional bias is as follows:

Definition I.B.11: Under the general mixed linear model, the total conditional bias (TCB) of $\underline{\hat{\mu}}$ as an estimator of $\underline{\mu}$ is defined as

$$TCB(\underline{\hat{\mu}},\underline{\mu}) = \sum_{i=1}^{n} E[\hat{\mu}_i - \mu_i \mid \underline{a}=\underline{\alpha}]. \tag{I.B.11}$$

Definition I.B.12: Under the general mixed linear model, the total mean squared error (TMSE) of $\underline{\hat{\mu}}$ as an estimator of $\underline{\mu}$ is given by

$$TMSE(\underline{\hat{\mu}},\underline{\mu}) = \sum_{i=1}^{n} E(\hat{\mu}_i - m_i)^2$$

$$= E[\underline{\hat{\mu}}-\underline{m}]'[\underline{\hat{\mu}}-\underline{m}]. \tag{I.B.12}$$

Definition I.B.13: Under the general mixed linear model, the total conditional mean squared error (TCMSE) of $\underline{\hat{\mu}}$ as an estimator of $\underline{\mu}$ is given by

$$TCMSE(\underline{\hat{\mu}},\underline{\mu}) = \sum_{i=1}^{n} E[(\hat{\mu}_i - \mu_i)^2 \mid \underline{a}=\underline{\alpha}]$$

$$= E\{[\underline{\hat{\mu}}-\underline{\mu}]'[\underline{\hat{\mu}}-\underline{\mu}] \mid \underline{a} = \underline{\alpha}\}. \tag{I.B.13}$$

The TMSE and TCMSE are sums of the corresponding mean squared errors of the individual components.

Let $\tilde{\underline{\beta}}$ and $\tilde{\underline{\alpha}}$ represent any solution to the normal equations (I.B.2), and $\hat{\underline{\beta}}$ and $\hat{\underline{\alpha}}$ any solution to the mixed-model equations (I.B.3). Let $\tilde{\underline{\mu}} = \underline{X\tilde{\beta}} + \underline{Z\tilde{\alpha}}$ and $\hat{\underline{\mu}} = \underline{X\hat{\beta}} + \underline{Z\hat{\alpha}}$. The Extended Gauss-Markov theorem assures us that, among all estimators of $\underline{\mu}$ that are linear, and componentwise unbiased, $\hat{\underline{\mu}}$ uniformly minimizes TMSE. An analogous result holds for $\hat{\underline{\mu}}$ and TCMSE. However, it is well-known that, if rank $(\underline{X}, \underline{Z}) \geq 3$, there exist nonlinear biased estimators that dominate $\tilde{\underline{\mu}}$ under the TCMSE criterion (see e.g., Arnold, 1981, Chapter 11).

Many of the unconditional properties and definitions described above required the matrix $\underline{D}$ to be known. In practice, at least some elements of $\underline{D}$ will generally be unknown. The traditional advice on estimating linear combinations of the elements of $\underline{\beta}$ and $\underline{\alpha}$ when $\underline{D}$ is unknown has been to first estimate $\underline{D}$ and to then proceed as though that estimate were the true matrix $\underline{D}$. Harville (1977) gave a review (primarily in the context of variance-component estimation) of various techniques for estimating $\underline{D}$ and σ_e^2. In general, different estimators for $\underline{D}$ will, upon their substitution in (I.B.5) and (I.B.6), produce different estimators for linear combinations of the elements of $\underline{\alpha}$ and $\underline{\beta}$. The conditional and unconditional mean squared errors of the latter estimators tend not to have simple expressions, making comparisons among these estimators difficult. We attempt to evaluate and compare the conditional and unconditional mean squared errors of these estimators in the context of one of the simplest versions of the general mixed linear model, namely, the balanced one-way random model. Even for this relatively simple model, the problem of evaluating the mean squared errors of the various estimators of linear combinations of fixed and random effects is by no means trivial.

C. Overview of the Remaining Chapters

In Chapter II, we review some basic, but very important, results for the balanced one-way classification. Essentially, we give these the general results of section I.B, as applied to this simple model. In section II.A, we discuss the fixed-effects model. In section II.B, we apply the mixed-model techniques discussed in Chapter I to the one-way random-effects model, taking the ratio $\rho = \sigma_e^2/(\sigma_e^2+J\sigma_a^2)$ to be known (and taking σ_e^2 and σ_a^2 to be as defined in section I.A). Section II.C relates the random model to a Bayesian formulation of the fixed model. By exploiting this relationship, various estimators of group contrasts under the random model can be interpreted as Bayes estimators of the same contrasts under the fixed model.

Chapter III is a key chapter. It discusses estimators of the ratio ρ and the corresponding estimators of linear combinations of realized values of random effects. A list of 18 estimators of ρ was compiled, based on the literature on variance component estimation. These estimators are separated into five categories and are summarized in section III.F.

Different estimators of ρ produce different estimators of the realized values of the random effects. In Chapters IV and V, we attempt to evaluate and compare the unconditional and conditional biases and mean squared errors of the latter estimators. In Chapter IV, we study the unconditional properties of the estimators, while Chapter V covers the conditional properties. The results of Chapter V reveal how estimators derived under the random model behave when the effects are in fact fixed. It is found that, in general, these estimators tend to have good ensemble conditional properties and to dominate the ordinary least-squares estimator under the total conditional mean squared error criterion. Convenient expressions are given in these two chapters for the biases and mean squared errors of two types of estimators. Expressions are also given for the biases and mean squared errors of the other three types of estimators. The latter expressions are more complicated than those for the first two types, but should prove useful in numerical studies.

II. THE BALANCED ONE-WAY CLASSIFICATION

A. The Balanced One-Way Fixed-Effects Model

The balanced one-way fixed model was defined in (I.A.2). We summarize some basic results about this model. A more detailed exposition is given, e.g., by Searle (1971, Chapter 4), and by Snedecor and Cochran (1980, Chapter 10).

As discussed in Chapter I, this model can be obtained from the balanced one-way random model (I.A.1) by conditioning on $a_i = \alpha_i$ $(i = 1,\ldots,I)$. It is a special case of the general mixed linear model (I.B.1) conditional on $\underline{a} = \underline{\alpha}$.

The parametric functions $\mu_1,\ldots,\mu_I$ are estimable and linearly independent. Every estimable function can be expressed as linear combinations of these functions. Thus, neither μ nor α_i is estimable. Note that $\mu^* = (1/I)\sum_h \mu_h$ and that $\alpha_i^* = \mu_i-\mu^*$, implying that both μ^* and α_i^* are estimable $(i = 1,\ldots,I)$.

The customary analysis-of-variance table for the balanced one-way fixed model is

Source	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Among	I-1	SS_a	MS_a	$\sigma_e^2+Js_\alpha^2$
Within	I(J-1)	SS_e	MS_e	σ_e^2
TOTAL	IJ-1			

(II.A.1)

The normal equations (I.B.2) become

$$\begin{aligned} IJ\tilde{\mu} + J\sum_h \tilde{\alpha}_h &= IJ\bar{y}_{..} \\ J\tilde{\mu} + J\tilde{\alpha}_i &= J\bar{y}_{i.} \quad (i=1,\ldots,I) \end{aligned} \tag{II.A.2}$$

or, in terms of estimable parameters:

$$J\tilde{\mu}_i = J\bar{y}_{i.} \qquad (i = 1,\ldots,I). \tag{II.A.3}$$

By solving the normal equations, we obtain the following best linear unbiased estimators (BLUE's) of the estimable parameters μ_i, α_i^*, and μ^*:

$$\tilde{\mu}_i = \bar{y}_{i.}, \quad \tilde{\alpha}_i^* = \bar{y}_{i.}-\bar{y}_{..}, \quad \text{and} \quad \tilde{\mu}^* = \bar{y}_{..} \ . \tag{II.A.4}$$

In the context of the balanced one-way random model, the estimators (II.A.4) are best linear conditionally unbiased estimators (BLCUE's). The conditional mean squared errors (CMSE's) of these estimators are:

$$\text{CMSE}(\tilde{\mu}_i,\mu_i) = \frac{\sigma_e^2}{J} \quad (i = 1,\ldots,I), \tag{II.A.5}$$

$$\text{CMSE}(\tilde{\alpha}_i^*,\alpha_i^*) = \left(\frac{I-1}{IJ}\right)\sigma_e^2 \quad (i = 1,\ldots,I), \tag{II.A.6}$$

$$\text{CMSE}(\tilde{\mu},\mu^*) = \frac{\sigma_e^2}{IJ}, \tag{II.A.7}$$

and

$$\text{CMSE}\left(\sum_h \ell_h\tilde{\mu}_h, \sum_h \ell_h\mu_h\right) = \frac{\sigma_e^2}{J}\left(\sum_h \ell_h^2\right), \tag{II.A.8}$$

where $\sum_h \ell_h \mu_h$ represents an arbitrary linear combination of $\mu_1, \ldots, \mu_I$.

Defining $\underline{\tilde{\mu}} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_I)'$, then the total conditional mean squared error (TCMSE) of $\underline{\tilde{\mu}}$ as an estimator of $\underline{\mu}$, as defined by (I.B.13), is

$$\text{TCMSE}(\underline{\tilde{\mu}}, \underline{\mu}) = J \sum_{i=1}^{I} E(\hat{\mu}_i - \mu_i)^2 = I\sigma_e^2. \qquad \text{(II.A.9)}$$

The total conditional bias (TCB) of $\underline{\tilde{\mu}}$, as defined by (I.B.11) is obviously zero, since each component of $\underline{\tilde{\mu}}$ is an unbiased estimate of the corresponding element of $\underline{\mu}$.

B. The Balanced One-Way Random-Effects Model

The balanced one-way random model was introduced in section I.A. In this section, we apply the definitions and results given in section I.B for the general mixed linear model to the one-way random model.

All linear combinations of μ and the α_i's are (unconditionally) estimable. To see this, note that $\ell_0 \bar{y}_{..}$ is a linear (unconditionally) unbiased estimator of the arbitrary parametric function $\tau = \ell_0 \mu + \sum_i \ell_i \alpha_i$.

For the balanced one-way random model, the matrix $\underline{D}$ defined in section I.B, reduces to $(\sigma_a^2/\sigma_e^2)\underline{I}$. Thus, the elements of $\underline{D}$ are known if and only if the variance ratio σ_a^2/σ_e^2 is known, or, equivalently, if the ratio of expected mean squares $\rho = \gamma_e/\gamma_a = \sigma_e^2/(\sigma_e^2+J\sigma_a^2)$ is known.

The analysis-of-variance table for the random model is the same as that for the fixed model, except for the expected value of the among-groups mean square. For the random model, this expected mean square is

$$E(MS)_a = \gamma_a = \sigma_e^2 + J\sigma_a^2.$$

In terms of the parametrization μ, σ_a^2, and σ_e^2, the parameter space of the random model is the collection of values that satisfy the restrictions

$$-\infty<\mu<\infty, \quad \sigma_e^2 > 0, \quad \sigma_a^2 \geq 0. \tag{II.B.1}$$

The random model can also be parametrized in terms of μ, γ_a, and γ_e or in terms of μ, ρ, and σ_e^2, and the restrictions on the parameter can be re-expressed as

$$-\infty<\mu<\infty, \quad \gamma_a \geq \gamma_e > 0, \tag{II.B.2}$$

or

$$-\infty<\mu<\infty, \quad \sigma_e^2 > 0, \quad 0<\rho\leq 1. \tag{II.B.3}$$

The statistics $\bar{y}_{..}$, SS_a, and SS_e form a set of complete sufficient statistics for μ, σ_a^2, and σ_e^2 (e.g., Graybill, 1976, Chapter 15). These three statistics are distributed independently as:

$$\bar{y}_{..} \sim N(\mu, \frac{\gamma_a}{IJ}) \tag{II.B.4}$$

$$SS_e \sim \gamma_e \chi^2(IJ-I) \tag{II.B.5}$$

$$SS_a \sim \gamma_a \chi^2(I-1) \tag{II.B.6}$$

where $\chi^2(n)$ denotes a central chi-square distribution with n degrees of freedom. Conditionally on $\underline{a}=\underline{\alpha}$, $\bar{y}_{..}$, SS_a and SS_e are distributed independently as

$$\bar{y}_{..} \sim N(\mu+\bar{\alpha}_{.}, \frac{\gamma_e}{IJ}) \tag{II.B.7}$$

$$SS_e \sim \sigma_e^2 \chi^2(IJ-I) \tag{II.B.8}$$

$$SS_a \sim \sigma_e^2 \chi^2(I-1,\lambda) \tag{II.B.9}$$

where λ is given by (I.A.14), and $\chi^2(n,\lambda)$ denotes a noncentral chi-square distribution with n degrees of freedom and noncentrality parameter λ.

For each i and j, the three statistics $\bar{y}_{..}$, $\bar{y}_{i.}-\bar{y}_{..}$, and $\bar{y}_{ij}-\bar{y}_{i.}$ are distributed independently, both unconditionally and conditionally on $\underline{a}=\underline{\alpha}$. To see this, note that their joint distribution is multivariate normal and that they are uncorrelated.

The elements of the variance-covariance matrix $\underline{V}$ of the vector of responses $\underline{y}$ are

$$\mathrm{Cov}(y_{ij}, y_{i'j'}) = \sigma_a^2 + \sigma_e^2, \quad \text{if } i=i' \text{ and } j=j',$$
$$= \sigma_a^2, \quad \text{if } i=i' \text{ and } j \neq j',$$
$$= 0, \quad \text{if } i \neq i'$$
$$(i,i'=1,\ldots,I;\ j,j'=1,\ldots,I). \quad \text{(II.B.10)}$$

If ρ is known and different from one, the mixed-model equations (I.B.3) become

$$IJ\hat{\hat{\mu}} + J \sum_h \hat{\hat{\alpha}}_h = IJ\bar{y}_{..}$$
$$J\hat{\hat{\mu}} + \frac{J\hat{\hat{\alpha}}_i}{(1-\rho)} = J\bar{y}_{i.} \quad (i = 1,\ldots,I). \qquad \text{(II.B.11)}$$

These equations are the same as the normal equations (II.A.2), except for the coefficient of $\hat{\hat{\alpha}}_i$ in the last I equations.

By solving system (II.B.11), we find that, if ρ is known and different from one, the (unconditional) best linear unbiased estimators (BLUE's) of μ_i, α_i, and μ are

$$\hat{\hat{\mu}}_i = (1-\rho)(\bar{y}_{i.} - \bar{y}_{..}) + \bar{y}_{..}$$
$$\hat{\hat{\alpha}}_i = (1-\rho)(\bar{y}_{i.} - \bar{y}_{..}) \qquad \text{(II.B.12)}$$
$$\hat{\mu} = \bar{y}_{..}$$

It is easy to verify that formulas (II.B.12) give the BLUE's of μ_i, α_i, and μ, for the case $\rho=1$ as well.

The limits of estimators (II.B.12) as ρ goes to zero give the corresponding estimators (II.A.4) for the fixed-effects model.

Note that estimator $\hat{\hat{\mu}}_i$ in (II.B.12) can be re-expressed as

$$\hat{\hat{\mu}}_i = (1-\rho)\bar{y}_{i.}+\rho\bar{y}_{..}, \tag{II.B.13}$$

i.e., as a convex linear combination of the sample mean of the ith group and the overall sample mean. Thus, the BLUE of μ_i is obtained by shrinking $\bar{y}_{i.}$, which is the BLCUE of μ_i, towards $\bar{y}_{..}$. For this reason, estimators of the form (II.B.13) are sometimes referred to as "shrinkage" estimators or "shrunken" estimators.

Let

$$\hat{\hat{\tau}} = \ell_0\hat{\hat{\mu}} + \sum_h \ell_h\hat{\hat{\alpha}}_h \tag{II.B.14}$$

represent the BLUE of the arbitrary parametric function τ, for the case ρ known.

Theorem II.B.1: If ρ is known, then the (unconditional) MSE of $\hat{\hat{\tau}}$ as an estimator of τ is given by

$$\begin{aligned} \text{MSE}(\hat{\hat{\tau}},\tau) &= E(\hat{\hat{\tau}}-t)^2 \\ &= \frac{\sigma_e^2}{IJ}\frac{1}{\rho}\{\ell_0^2 + (1-\rho)[1+\rho(I-1)]\sum_i \ell_i^2 \\ &\quad + (1-\rho)^2 \sum_i \sum_{i'\neq i} \ell_i\ell_{i'} - 2\ell_0(1-\rho)\sum_i \ell_i\}. \end{aligned}$$

For the case of $\rho\neq 1$, Theorem II.B.1 is a special case of Theorem I.B.4. For the case $\rho=1$, the proof of Theorem II.B.1 is straightforward. In particular, we obtain the following MSEs for estimators (II.B.12):

$$\mathrm{MSE}(\hat{\hat{\mu}}_i,\mu_i) = \frac{\sigma_e^2}{IJ}\frac{1}{\rho}\{1+(1-\rho)[1+\rho(I-1)]\}, \tag{II.B.15}$$

$$\mathrm{MSE}(\hat{\hat{\alpha}}_i,\alpha_i) = \frac{\sigma_e^2}{IJ}\frac{(1-\rho)}{\rho}[1+\rho(I-1)], \tag{II.B.16}$$

and

$$\mathrm{MSE}(\hat{\hat{\mu}},\mu) = \frac{\sigma_e^2}{IJ}\frac{1}{\rho}. \tag{II.B.17}$$

Let

$$\underline{\hat{\hat{\mu}}} = (\hat{\hat{\mu}}_1,\ldots,\hat{\hat{\mu}}_I)' \tag{II.B.18}$$

Expression (II.B.15) does not depend on i, and therefore

$$\begin{aligned}\mathrm{TMSE}(\underline{\hat{\hat{\mu}}},\underline{\mu}) &= \sum_i\sum_j \mathrm{MSE}(\hat{\hat{\mu}}_i,\mu_i)\\ &= \frac{\sigma_e^2}{\rho}\{1+(1-\rho)[1+\rho(I-1)]\}.\end{aligned} \tag{II.B.19}$$

Obviously, the total bias of $\underline{\hat{\hat{\mu}}}$, as defined in (I.B.10), is zero, since the $\hat{\hat{\mu}}_i$'s estimate the μ_i's unbiasedly.

We have been considering the case where ρ is known. In subsequent chapters, we consider estimators of μ_i for the case when ρ is unknown. More specifically, we consider estimators of the form

$$\hat{\mu}_i = (1-\hat{\rho})(\bar{y}_{i.}-\bar{y}_{..}) + \bar{y}_{..}$$

$$= (1-\hat{\rho})\bar{y}_{i.} + \hat{\rho}\bar{y}_{..}$$

$$= \bar{y}_{i.} - \hat{\rho}(\bar{y}_{i.}-\bar{y}_{..}), \qquad \text{(II.B.20)}$$

where $\hat{\rho}$ is a function of SS_e and SS_a, that can be regarded as an estimator of ρ. Note that this estimator, like that given by (II.B.13), is a shrinkage estimator. The corresponding estimators of α_i, μ, and τ are of the form

$$\hat{\alpha}_i = (1-\hat{\rho})(\bar{y}_{i.}-\bar{y}_{..}), \qquad \text{(II.B.21)}$$

$$\hat{\mu} = \hat{\hat{\mu}} = \bar{y}_{..}, \qquad \text{(II.B.22)}$$

and

$$\hat{\tau} = \ell_0\hat{\mu} + \sum_h \ell_h\hat{\alpha}_h, \qquad \text{(II.B.23)}$$

respectively. Estimators (II.B.20)-(II.B.23) have an empirical Bayes interpretation, as discussed in the following section.

C. The Random-Effects Model as a Bayesian Formulation of the Fixed-Effects Model

The balanced one-way fixed model can be re-expressed as

$$y_{ij} = m_i + e_{ij} \quad (i = 1,\ldots,I; \quad j = 1,\ldots,J), \qquad \text{(II.C.1)}$$

where $m_1,\ldots,m_I$ are unknown parameters and $e_{11},e_{12},\ldots,e_{IJ}$ are identically and independently distributed as $N(0,\sigma_e^2)$.

In Bayesian inference, it is assumed that the experimenter can express his prior knowledge about $m_1,,,,.m_I$ in the form of a (prior) probability distribution. Let us suppose that, a priori, $m_1,\ldots,m_I$ are identically and independently distributed as $N(\mu,\sigma_a^2)$. Let $a_i = m_i - \mu (i = 1,\ldots,I)$. The model (II.C.1) and the prior assumptions can be summarized as:

$$y_{ij} = \mu + a_i + e_{ij} \quad (i = 1,\ldots,I; \; j = 1,\ldots,J), \qquad \text{(II.C.2)}$$

where $a_1,\ldots,a_I$ are identically distributed as $N(0,\sigma_a^2)$, $e_{11},e_{12},\ldots,e_{IJ}$ are identically distributed as $N(0,\sigma_e^2)$, and $a_1,\ldots,a_I,e_{11},e_{12},\ldots,e_{IJ}$ are statistically independent.

This model is essentially the same as the balanced one-way random model (I.A.1). Thus, we can interpret the random-effects model (I.A.1) as a particular Bayesian formulation of the fixed-effects model (II.C.1).

The likelihood function for model (II.C.1) is

$$\ell(\underline{m},\sigma_e^2|\underline{y}) \propto (\sigma_e^2)^{\frac{-IJ}{2}} \cdot \exp\{-\frac{1}{2} \sum_i \sum_j \frac{(y_{ij}-m_i)^2}{\sigma_e^2}\}$$

$$= (\sigma_e^2)^{-\frac{I(J-1)}{2}} (\sigma_e^2)^{-\frac{I}{2}} \exp\{-\frac{1}{2} \frac{1}{\sigma_e^2} [SS_e$$

$$+ J \sum_i (\bar{y}_{i.}-m_i)^2]\} \qquad \text{(II.C.3)}$$

The joint distribution of $\underline{m}$ and $\underline{y}$ conditional on μ, σ_a^2, and σ_e^2, is given by

$$p(\underline{m},\underline{y}|\mu,\sigma_a^2,\sigma_e^2) \propto \ell(\underline{m},\sigma_e^2|\underline{y})$$

$$\cdot\ (\sigma_a^2)^{-\frac{I}{2}} \exp\{-\frac{1}{2} \sum_i \frac{(m_i-\mu)^2}{\sigma_a^2}\}. \qquad \text{(II.C.4)}$$

After some algebraic manipulations, we find that, conditionally on $\underline{y}$, μ, σ_a^2, and σ_e^2, the quantities $m_1,\ldots,m_I$ are distributed independently with

$$m_i \sim N[(1-\rho)\bar{y}_{i.}+\rho\mu,\ \frac{\sigma_e^2}{J}(1-\rho)] \quad (i = 1,\ldots,I). \qquad \text{(II.C.5)}$$

Therefore,

$$E(m_i|\underline{y},\mu,\sigma_a^2,\sigma_e^2) = E(m_i|\underline{y},\mu,\rho,\sigma_e^2)$$

$$= (1-\rho)\bar{y}_{i.}+ \rho\mu.$$

Since this expression does not depend on σ_e^2, we have that

$$E(m_i|\underline{y},\mu,\rho) = (1-\rho)\bar{y}_{i.}+\rho\mu \qquad \text{(II.C.6)}$$

Similarly,

$$E(a_i|\underline{y},\mu,\rho) = (1-\rho)(\bar{y}_i-\mu), \qquad \text{(II.C.7)}$$

and

$$E(\ell_0\mu + \sum_i \ell_i a_i|\underline{y},\mu,\rho)$$

$$= \ell_0\mu + \sum_i \ell_i(1-\rho)(\bar{y}_{i.}-\mu). \qquad \text{(II.C.8)}$$

Under the quadratic loss, the Bayes estimator of a parametric function equals its posterior mean. Subsequently, we use Bayes estimator synonymously with posterior mean. A full-fledged Bayesian approach would require that the prior distributions be completely specified, i.e., that μ, σ_a^2 and σ_e^2 be taken to be known or that prior distributions be put on them as well. If μ and ρ are known, the Bayes estimator of m_i is

$$\hat{m}_i = (1-\rho)(\bar{y}_{i.}-\mu)+\mu, \tag{II.C.9}$$

and expression (II.C.9) can be used to obtain Bayes estimators of other functions.

If μ and ρ are unknown, a Bayesian statistician needs to specify a joint prior distribution for these quantities to complete the specification of his prior distribution. We have that

$$\begin{aligned} E(m_i|\underline{y}) &= E[(1-\rho)\bar{y}_{i.}+\rho\mu|\underline{y}] \\ &= [1-E(\rho|\underline{y})]\bar{y}_{i.} + E(\rho\mu|\underline{y}), \end{aligned} \tag{II.C.10}$$

which gives the general form of the Bayes estimator of m_i.

An alternative approach would be to replace μ and ρ in (II.C.9) by "reasonable" estimators of these parameters (not necessarily equal to their posterior means). This approach is known as the empirical Bayes approach. For the balanced one-way random model, the BLUE of μ is equal to $\bar{y}_{..}$, which suggests that we consider estimators of m_i of the form

$$\hat{m}_i = (1-\hat{\rho})\bar{y}_{i.} + \hat{\rho}\bar{y}_{..}, \tag{II.C.11}$$

where $\hat{\rho}$ is some estimator of ρ. Note that these estimators have the same form as the estimators (II.B.20) of μ_i of the previous section. For this reason, estimators of the form (II.C.11) can be regarded either as estimators of μ_i (the realized value of m_i) under the random model (I.A.1), or as empirical Bayes estimators of m_i under the fixed model (II.C.1).

III. ESTIMATION OF THE EXPECTED-MEAN-SQUARE RATIO

A. Introduction

Take the model to be the one-way random model. The problem of estimating the ratio $\rho = \gamma_e/\gamma_a$ of expected mean squares is closely related to that of estimating the variance components σ_e^2 and σ_a^2. The latter problem has received considerable attention in the literature. In this chapter, we shall review some of the techniques that have been proposed for variance component estimation and obtain the corresponding estimators of the expected-mean-square ratio.

The parameter ρ is restricted to the interval $0<\rho\leq 1$, however, some "estimators" of ρ can exceed one with nonzero probability. Let $\hat{\rho}$ denote an arbitrary estimator of ρ. If $\hat{\rho}>1$, the estimator (II.B.20) of μ_i "shrinks" $\bar{y}_{i.}$ past $\bar{y}_{..}$. With this "overshrinking", the ranking of the $\hat{\mu}_i$'s is in reverse order from that of the $\bar{y}_{i.}$'s; the smallest $\bar{y}_{i.}$ produces the largest $\hat{\mu}_i$, which seems undesirable. However, despite these shortcomings, estimators that allow overshrinkage are often considered in the literature. The properties of these estimators are generally much easier to derive than those of estimators that preserve the ordering of the $\bar{y}_{i.}$'s.

B. Analysis of Variance Estimators

These estimators are obtained by equating mean squares in the analysis of variance (ANOVA) table with their respective expectations and then solving the resulting system of equations. We have that

$$E(MS_a) = \gamma_a = \sigma_e^2 + J\sigma_a^2 \tag{III.B.1}$$

and

$$E(MS_e) = \gamma_e = \sigma_e^2. \tag{III.B.2}$$

Thus, the ANOVA estimators of γ_a and γ_e are

$$\hat{\gamma}_e = MS_e, \quad \hat{\gamma}_a = MS_a, \tag{III.B.3}$$

and the corresponding estimator of ρ (labelled E1 for convenience) is

$$\text{E1:}\ \hat{\rho} = \frac{MS_e}{MS_a} = \frac{SS_e}{SS_a}\,\frac{(I-1)}{I(J-1)}\ . \tag{III.B.4}$$

The ANOVA technique produces unbiased estimators of expected mean squares and variance components. However, these estimators can assume values outside the parameter space. To avoid an estimate of ρ greater than one, the estimator $\hat{\rho}$ given by (III.B.3) can be truncated, producing the estimator

$$\text{E2:}\ \hat{\rho} = \min\{\frac{SS_e}{SS_a}\,\frac{(I-1)}{I(J-1)},\ 1\}. \tag{III.B.5}$$

The untruncated version, E1, is the proper ANOVA estimator of ρ. The truncated version, E2, will be called the truncated ANOVA estimator.

C. Maximum Likelihood and Related Estimators

The following presentation is based on that of Harville (1978).

We consider the estimation of the ratio of expected mean squares by maximum likelihood (ML), by Patterson and Thompson's (1971, 1974) restricted or modified maximum likelihood (REML), and by a pseudo-Bayesian modification of REML described by Harville (1977). In deriving maximum likelihood estimators, we take advantage of the results of Chapter II. Recall that $\bar{y}_{..}$, SS_e, and SS_a form a complete set of sufficient statistics for μ, γ_e, and γ_a. These three statistics are jointly independent, with marginal distributions given by (II.B.4), (II.B.5), and (II.B.6). Thus, the ML estimates of μ, γ_e, and γ_a are given by the values of these parameters that maximize the function

$$\gamma_a^{-1/2} \cdot (SS_e/\gamma_e)^{\frac{I(J-1)}{2}-1} \cdot (SS_a/\gamma_a)^{\frac{I-1}{2}-1} \cdot (\gamma_e)^{-1} \cdot (\gamma_a)^{-1}$$

$$\cdot \exp\{-(\tfrac{1}{2})[(IJ/\gamma_a)(\bar{y}_{..}-\mu)^2 + SS_e/\gamma_e + SS_a/\gamma_a]\},$$

subject to the restrictions $\gamma_a \geq \gamma_e > 0$.

Clearly, the ML estimate of μ is $\bar{y}_{..}$, and the ML estimators of γ_a and γ_e are those values of γ_a and γ_e that maximize the function

$$L_M(\gamma_e,\gamma_a) = -\frac{1}{2}\{I \log(\gamma_a) + I(J-1)\log(\gamma_e) + \frac{SS_a}{\gamma_a} + \frac{SS_e}{\gamma_e}\}, \tag{III.C.1}$$

subject to the restrictions

$$\gamma_a \geq \gamma_e > 0. \tag{III.C.2}$$

If we subtract $\bar{y}_{..}$, the ML estimate of μ, from the response y_{ij} ($i = 1,\ldots,I$; $j = 1,\ldots,J$) we obtain a set of error contrasts

$$\hat{e}_{ij} = y_{ij}-\bar{y}_{..} \quad (i = 1,\ldots,I; \quad j = 1,\ldots,J). \tag{III.C.3}$$

(By definition, an error contrast is a linear unbiased estimator of zero.) Since $\sum_i \sum_j \hat{e}_{ij} = 0$, only IJ-1 of the error contrasts (III.C.3) are linearly independent. If we proceed as though the data consists of any IJ-1 of the $\hat{e}_{ij}$'s, then SS_a and SS_e are complete sufficient statistics for γ_a and γ_e. The logarithm of the resulting likelihood function is (except for an additive constant),

$$L_R(\gamma_e,\gamma_a) = -\frac{1}{2}\{(I-1)\log(\gamma_a) + I(J-1)\log(\gamma_e) + SS_a/\gamma_a + SS_e/\gamma_e\}. \tag{III.C.4}$$

The REML estimator of γ_e and γ_a are defined to be those values of these parameters that maximize L_R subject to the restrictions (III.C.2).

Harville (1977) proposed the following modification of the REML approach: take as estimates of γ_e and γ_a those values obtained by maximizing (over the parameter space) the product of the likelihood function of the error contrasts and the Jeffreys' noninformative prior distribution derived from that likelihood function. It follows from the results of Box and Tiao (1973, Section 5.2), that the Jeffreys' prior distribution is the distribution with "p.d.f."

$$p(\gamma_e,\gamma_a) = (\gamma_e\gamma_a)^{-1} \quad (\gamma_a \geq \gamma_e > 0). \tag{III.C.5}$$

Therefore, the estimators of γ_a and γ_e are obtained by maximizing the function

$$L_J(\gamma_e,\gamma_a) = -\tfrac{1}{2}\{(I+1)\log(\gamma_a) + (IJ-I+2)\log(\gamma_e) + SS_a/\gamma_a + SS_e/\gamma_e\} \tag{III.C.6}$$

subject to the restrictions (III.C.2).

Harville (1977, 1978) referred to this approach as a pseudo-Bayesian modification of REML. It is equivalent to obtaining the mode of a particular joint posterior distribution of γ_e and γ_a.

The three functions L_M, L_R, and L_J are all of the form

$$L = -\frac{1}{2}\{(f_a+k_a)\log(\gamma_a) + (f_e+k_e)\log(\gamma_e) + \frac{SS_a}{\gamma_a} + \frac{SS_e}{\gamma_e}\},$$

(III.C.7)

where f_a and f_e are the degrees of freedom (I-1) and I(J-1), respectively. The choices for k_a and k_e that give L_M, L_R, and L_J are

$$L_M: \quad k_a = 1, \; k_e = 0$$

$$L_R: \quad k_a = k_e = 0 \qquad \text{(III.C.8)}$$

$$L_J: \quad k_a = k_e = 2$$

For $\gamma_a>0$ and $\gamma_e>0$, we have

$$\frac{\partial L}{\partial \gamma_j} = -(\tfrac{1}{2})(f_j+k_j)/\gamma_j + (\tfrac{1}{2})SS_j/\gamma_j^2 \quad (j = a,e). \qquad \text{(III.C.9)}$$

The equation (III.C.9) = 0 has as its solution

$$\gamma_j = W_j = SS_j/(f_j+k_j) \quad (j = a,e). \qquad \text{(III.C.10)}$$

It is easy to check that the solution (III.C.10) maximizes L for $\gamma_a>0$ and $\gamma_e>0$, but the solution can violate the constraint $\gamma_a \geq \gamma_e$. If we ignored the restriction $\gamma_a \geq \gamma_e$, we would obtain the following estimators

$$\text{ML:} \quad \hat{\gamma}_a = SS_a/I, \quad \hat{\gamma}_e = SS_e/(IJ-I) \qquad \text{(III.C.11)}$$

$$\text{REML:} \quad \hat{\gamma}_a = SS_a/(I-1), \quad \hat{\gamma}_e = SS_e/(IJ-I) \qquad \text{(III.C.12)}$$

modified REML: $\hat{\gamma}_a = SS_a/(I+1)$, $\hat{\gamma}_e = SS_e/(IJ-I+2)$ (III.C.13)

The corresponding estimators of $\rho = \gamma_e/\gamma_a$ would be

$$\text{E3: } \hat{\rho} = \frac{SS_e}{SS_a}\,\frac{1}{(J-1)} \qquad \text{(III.C.14)}$$

$$\text{E4: } \hat{\rho} = \frac{SS_e}{SS_a}\,\frac{(I-1)}{(IJ-I)} \qquad \text{(III.C.15)}$$

$$\text{E5: } \hat{\rho} = \frac{SS_e}{SS_a}\,\frac{(I+1)}{(IJ-I+2)} \qquad \text{(III.C.16)}$$

Note that estimators (III.C.12) of γ_a and γ_e are equal to the ANOVA estimators (III.B.3) and, hence, are unbiased. [They are, in fact, the minimum variance quadratic translation-invariant unbiased estimators, as discussed by Harville (1977) and LaMotte (1970, 1971, and 1973).] They produce an estimator of ρ (E4) that is identical to E1. The modified-REML estimators of γ_a and γ_e have an interesting property: among all estimators of γ_j of the form $c_j SS_j$ ($j = a,e$), where c_j is a constant, estimators (III.C.13) have uniformly smallest mean squared error.

Estimators (III.C.11), (III.C.12), and (III.C.13) do not take into account the restrictions on the parameter space and, therefore, they are not the true ML, REML, or modified-REML estimators. It can be shown that the values of γ_a and γ_e that maximize L subject to the restrictions $\gamma_a \geq \gamma_e > 0$ are

$$\hat{\gamma}_a = W_a \quad \text{if} \quad W_a \geq W_e$$

$$= W_{ae} \quad \text{if} \quad W_a < W_e \qquad \text{(III.C.17)}$$

and

$$\hat{\gamma}_e = W_e \quad \text{if} \quad W_a \geq W_e$$

$$= W_{ae} \quad \text{if} \quad W_a < W_e \ , \qquad \text{(III.C.18)}$$

where $W_{ae} = (h_a W_a + h_e W_e)/(h_a + h_e)$ and $h_j = f_j + k_j$ $(j = a,e)$. These results are particular cases of results given by Harville (1978). The estimators of γ_a and γ_e, given by (III.C.17) and (III.C.18), lead to the following estimator of ρ:

$$\hat{\rho} = \min\{\frac{W_e}{W_a}, 1\}. \qquad \text{(III.C.19)}$$

Thus, the true ML, REML, and modified-REML estimators of ρ are

$$\text{E6: } \hat{\rho} = \min\{\frac{SS_e}{SS_a} \frac{1}{(J-1)}, 1\}, \qquad \text{(III.C.20)}$$

$$\text{E7: } \hat{\rho} = \min\{\frac{SS_e}{SS_a} \frac{(I-1)}{(IJ-I)}, 1\}, \qquad \text{(III.C.21)}$$

and

$$\text{E8: } \hat{\rho} = \min\{\frac{SS_e}{SS_a} \frac{(I+1)}{(IJ-I+2)}, 1\}, \qquad \text{(III.C.22)}$$

respectively. Note that these are truncated versions of estimators E3, E4, and E5.

D. Bayes Estimation Using Jeffreys' Noninformative Prior

Jeffreys (1961) proposed a rule to obtain noninformative priors. For a single parameter, his prior "p.d.f." is proportional to the square root of Fisher's measure of information. For multivariate problems, his prior "p.d.f." for the parameter vector is proportional to the square root of the determinant of the information matrix. Jeffreys' rule is invariant under parametric transformations.

Box and Tiao (1973, Chapters 1 and 5) give a detailed description of Jeffreys' approach and its application to the balanced one-way random model. They recommend a prior distribution in which the location parameter μ is taken to be distributed independently of the expected mean squares γ_e and γ_a. Using Jeffreys' rule, they arrive at a noninformative prior distribution in which μ, $\log(\gamma_a)$, and $\log(\gamma_e)$ are statistically independent with locally uniform distributions. Thus, the noninformative prior distribution has the "p.d.f."

$$p(\mu,\gamma_a,\gamma_e) = p_1(\mu)p_2(\gamma_a,\gamma_e), \qquad \text{(III.D.1)}$$

with

$$p_1(\mu) \propto \text{a constant}$$

and

$$p_2(\gamma_a,\gamma_e) \propto (\gamma_a\gamma_e)^{-1}$$

$$(\gamma_a \geq \gamma_e > 0).$$

Alternatively, in terms of μ, σ_a^2, and σ_e^2, the prior "p.d.f." is

$$p(\mu,\sigma_a^2,\sigma_e^2) = p_1(\mu)p_3(\sigma_a^2,\sigma_e^2) \propto \sigma_e^{-2}(\sigma_e^2+J\sigma_a^2)^{-1} \qquad \text{(III.D.2)}$$

$$(\gamma_a \geq \gamma_e > 0).$$

The likelihood function is

$$\ell(\mu,\gamma_a,\gamma_e|\underline{y}) \propto \gamma_a^{-1/2}(SS_e/\gamma_e)^{\frac{I(J-1)}{2}-1}(SS_a/\gamma_a)^{\frac{I-1}{2}-1}$$

$$\cdot\ \gamma_e^{-1}\gamma_a^{-1} \exp\{-(\tfrac{1}{2})[(IJ/\gamma_a)(\bar{y}_{..}-\mu)^2$$

$$+ SS_e/\gamma_e + SS_a/\gamma_a]\}, \qquad \text{(III.D.3)}$$

as discussed in the previous section.

By combining this likelihood function with the prior (III.D.1), the posterior p.d.f. of (μ,γ_a,γ_e) is found to be

$$p_4(\mu,\gamma_a,\gamma_e|\underline{y}) \propto (\gamma_e)^{-(\frac{1}{2}f_e+1)}(\gamma_a)^{-(\frac{1}{2}f_a+1)-\frac{1}{2}}$$

$$\cdot\ \exp\{-(\tfrac{1}{2})[SS_e/\gamma_e + SS_a/\gamma_a + (IJ/\gamma_a)(\bar{y}_{..}-\mu)^2]\}$$

$$(-\infty<\mu<\infty,\quad \gamma_a\geq\gamma > 0), \qquad \text{(III.D.4)}$$

where $f_e = I(J-1)$ and $f_a = I-1$.

To obtain the marginal p.d.f. of (γ_a,γ_e), we integrate expression (III.D.4) with respect to μ, obtaining

$$p_5(\gamma_a,\gamma_e|\underline{y}) \propto (\gamma_e)^{-(\frac{1}{2}f_e+1)}(\gamma_a)^{-(\frac{1}{2}f_a+1)}$$

$$\cdot \exp[-(\tfrac{1}{2})(SS_a/\gamma_a + SSE/\gamma_e)]$$

$$(\gamma_a \geq \gamma_e > 0) \qquad \text{(III.D.5)}$$

Note that the posterior p.d.f. (III.D.5) is the same as that whose logarithm is given by (III.C.6) (except for an additive constant). Aside from the restrictions on the parameter space, the p.d.f. (III.D.5) is the product of two inverted gamma p.d.f.'s. Thus, if these restrictions are ignored, the marginal p.d.f. of ρ is

$$p_6(\rho|\underline{y}) = \frac{MS_a}{MS_e}\, p[F_{f_a,f_e} = \frac{MS_a}{MS_e}\rho]$$

$$(0<\rho<\infty), \qquad \text{(III.D.6)}$$

where

$$p[F_{n_1,n_2} = z] = \frac{(n_1/n_2)^{n_1/2} z^{(n_1-2)/2}}{B(\frac{n_1}{2},\frac{n_2}{2})[1+(n_1/n_2)z]^{(n_1+n_2)/2}}$$

$$(0<z<\infty) \qquad \text{(III.D.7)}$$

denote the density of an F distribution with n_1 and n_2 degrees of freedom evaluated at z. We shall refer to (III.D.6) as the untruncated marginal posterior p.d.f. of ρ.

If I>3, the mode of the distribution with this p.d.f. is

$$\text{E9:}\ \hat{\rho} = \frac{SS_e}{SS_a}\,\frac{(I-3)}{(IJ-I+2)}\ . \qquad \text{(III.D.8)}$$

If $I \leq 3$, the p.d.f. is J-shaped. If $I(J-1)>2$, the mean of the distribution with p.d.f. (III.D.6) is

$$\text{E10:}\ \hat{\rho} = \frac{SS_e}{SS_a}\,\frac{(I-1)}{(IJ-I-2)}\ . \qquad \text{(III.D.9)}$$

If $I(J-1) \leq 2$, the mean of this distribution is infinite.

It can be shown that, when the restrictions $\gamma_a \geq \gamma_e > 0$ are taken into account, the marginal posterior p.d.f. of ρ is

$$p_7(\rho|\underline{y}) = \frac{(MS_a/MS_e)\,p[F_{f_a,f_e} = (MS_a/MS_e)\rho]}{\Pr[F_{f_a,f_e} < MS_a/MS_e]}$$

$$(0<\rho\leq 1), \qquad \text{(III.D.10)}$$

where F_{n_1,n_2} denotes a random variable whose distribution is F with n_1 and n_2 degrees of freedom. We shall refer to (III.D.10) as the truncated marginal posterior p.d.f. of ρ. The mode of this distribution, for I>3, is the following truncated version of estimator E9:

$$\text{E11:}\ \hat{\rho} = \min\{\frac{SS_e}{SS_a}\,\frac{(I-3)}{(IJ-I+2)},\ 1\}. \qquad \text{(III.D.11)}$$

If $I(J-1)>2$, the mean of the truncated marginal posterior distribution is:

$$\text{E12: } \hat{\rho} = \frac{SS_e}{SS_a} \frac{(I-1)}{(IJ-I-2)} \cdot \frac{I_x(c+1,d-1)}{I_x(c,d)}, \tag{III.D.12}$$

where $x = SS_a/(SS_e+SS_a)$, $c = (I-1)/2$, $d = I(J-1)/2$,

$$I_x(c,d) = \frac{\int_0^x t^{c-1}(1-t)^{d-1}dt}{B(c,d)} \tag{III.D.13}$$

denotes the incomplete beta function ratio. If $I(J-1)\leq 2$, the mean of this distribution does not exist.

Estimator E12 is obviously contained in the interval (0,1], since it is the mean of a distribution that assigns probability 1 to this interval. In fact, an upper bound smaller than one can be obtained, as we now show. Estimator E12 can be rewritten as

$$\hat{\rho} = \frac{c}{(d-1)} \frac{(1-x)}{x} \frac{I_x(c+1,d-1)}{I_x(c,d)} \tag{III.D.14}$$

with c, d, and x as defined in (III.D.12). The hypergeometric function is defined as

$$F(\alpha,\beta;\gamma;x) = 1 + \frac{\alpha\beta}{\gamma}\frac{x}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{x^2}{2!} + \ldots \quad (\gamma>0)$$

(see, e.g., Patel et al., 1976, Sections 10.4 and 10.9). $F(.,.;\gamma;.)$ is obviously a decreasing function of γ. Now, using results from Sections 10.4 and 10.9 of Patel et al. (1976), we have that

$$\hat{\rho} = \frac{c}{(c+1)}(1-x)\,\frac{F(c+1,2-d;c+2;x)}{F(c,1-d;c+1;x)}$$

$$= \frac{c}{(c+1)}\,\frac{F(1,c+d;c+2;x)}{F(1,c+d;c+1;x)} < \frac{c}{(c+1)} = \frac{I-1}{I+1} < 1. \qquad \text{(III.D.15)}$$

Therefore, estimator E12 is strictly less than $(I-1)/(I+1)$ (for $I>1$ and $IJ-I>2$).

E. Proper Bayes Estimation

Expression (III.D.5) suggests that a convenient proper prior p.d.f. for γ_a and γ_e is

$$p_8(\gamma_a,\gamma_e) \propto (\gamma_e)^{-(\frac{1}{2}f_e^*+1)}(\gamma_a)^{-(\frac{1}{2}f_a^*+1)}$$

$$\cdot \exp[-(\tfrac{1}{2})(SS_a^*/\gamma_a + SS_e^*/\gamma_e)]$$

$$(\gamma_a \geq \gamma_e > 0), \qquad \text{(III.E.1)}$$

where f_e^*, f_a^*, SS_a^*, and SS_e^* are arbitrary positive constants. In deciding on values for f_e^*, f_a^*, SS_a^*, and SS_e^*, it may be useful to think of these quantities as the corresponding degrees of freedom and sums of squares from a previous data set that follows a one-way random model.

The likelihood of a set of linearly independent error contrasts is

$$\ell(\gamma_a,\gamma_e|SS_a,SS_e) \propto \gamma_a^{-1}\gamma_e^{-1}(SS_a/\gamma_a)^{1/2\ f_a-1}(SS_e/\gamma_e)^{1/2\ f_e-1}$$

$$\cdot \exp\{-(\tfrac{1}{2})(SS_a/\gamma_a + SS_e/\gamma_e)\}$$

$$(\gamma_a \geq \gamma_e > 0), \qquad \text{(III.E.2)}$$

as discussed in Section III.D.

Combining this likelihood function with the prior p.d.f. (III.E.1), we obtain the posterior p.d.f.

$$p_9(\gamma_a,\gamma_e|SS_a,SS_e) \propto (\gamma_a)^{-[(\frac{1}{2})(f_a+f_a^*)+1]}$$

$$\cdot (\gamma_e)^{-[(\frac{1}{2})(f_e+f_e^*)+1]}$$

$$\cdot \exp\{-(\tfrac{1}{2})[(SS_a+SS_a^*)/\gamma_a + (SS_e+SS_e^*)\gamma_e]\}$$

$$(\gamma_a \geq \gamma_e > 0). \qquad \text{(III.E.3)}$$

Note that this expression has the same general form as expression (III.D.5), with f_a in (III.D.5) being replaced by $f_a+f_a^*$ in (III.E.3), f_e by $f_e+f_e^*$, SS_a by $SS_a+SS_a^*$, and SS_e by $SS_e+SS_e^*$. Thus, it is possible to deduce the posterior mean and mode of ρ from the results of the previous section.

If $I + f_a^* > 3$, the mode of the untruncated marginal posterior distribution of ρ is

$$\text{E13:}\quad \hat{\rho} = \frac{(SS_e+SS_e^*)(I+f_a^*-3)}{(SS_a+SS_a^*)(IJ-I+f_e^*+2)}, \qquad \text{(III.E.4)}$$

and the mode of the truncated marginal posterior distribution is

$$\text{E14: } \hat{\rho} = \min\{\frac{(SS_e+SS_e^*)(I+f_a^*-3)}{(SS_a+SS_a^*)(IJ-I+f_e^*+2)},1\} \qquad \text{(III.E.5)}$$

If $I(J-1)+f_2^* > 2$, the mean of the untruncated marginal posterior distribution of ρ is

$$\text{E15: } \hat{\rho} = \frac{(SS_e+SS_e^*)(I+f_a^*-1)}{(SS_a+SS_a^*)(IJ-I+f_e^*-2)}, \qquad \text{(III.E.6)}$$

and the mean of the truncated marginal posterior distribution is

$$\text{E16: } \hat{\rho} = \frac{(1-x)}{x}\,\frac{c}{(d-1)}\,\frac{I_x(c+1,d-1)}{I_x(c,d)}, \qquad \text{(III.E.7)}$$

with

$$x = (SS_a+SS_a^*)/(SS_e+SS_e^*+SS_a+SS_a^*),$$

$$c = (1/2)(f_a+f_a^*) = (1/2)(I-1+f_a^*),$$

and

$$d = (1/2)(f_e+f_e^*) = (1/2)(IJ-I+f_e^*).$$

Estimators E13-E16 were derived by assuming the prior (III.E.1) and taking the data to be linearly independent error contrasts. This approach is equivalent (for purposes of inference about ρ) to taking the data to be the vector $\underline{y}$ and taking the prior "p.d.f." of μ, γ_a, and γ_e to be

$$p_{10}(\mu,\gamma_a,\gamma_e) = p_1(\mu)p_8(\gamma_a,\gamma_e), \qquad \text{(III.E.8)}$$

where

$$p_1(\mu) \propto \text{a constant}, \tag{III.E.9}$$

and $p_8(\gamma_a, \gamma_e)$ is given by (III.E.1). As Hill (1977) notes, the use of the prior of the general form (III.E.8), with $p_1(\mu)$ given by (III.E.9), is tantamount to basing inference about the variance components solely upon SS_e and SS_a.

If a proper prior distribution for μ is preferred, it is convenient to consider a prior distribution in which the marginal distribution of γ_a and γ_e is given by (III.E.1) and, conditional on γ_a and γ_e,

$$\mu \sim N(\bar{y}^*_{..}, \frac{\gamma_a}{N^*}), \tag{III.E.10}$$

where $\bar{y}^*_{..}$ is an arbitrary constant and $N^* = f^*_a + f^*_e + 1$.

Combining this prior with the likelihood (III.D.3), we obtain the posterior p.d.f.

$$p_{11}(\mu, \gamma_a, \gamma_e | \underline{y}) \propto (\gamma_a)^{-(\frac{1}{2})(f_a+f^*_a)-2} (\gamma_e)^{-(\frac{1}{2})(f_e+f^*_e)-1}$$

$$\cdot \exp\{-(\tfrac{1}{2})[(SS_a+SS^*_a)/\gamma_a + (SS_e+SS^*_e)/\gamma_e + A]\},$$

$$(-\infty<\mu<\infty, \quad \gamma_a \geq \gamma_e > 0), \tag{III.E.11}$$

where

$$A = (N/\gamma_a)(\bar{y}_{..}-\mu)^2 + (N^*/\gamma_a)(\bar{y}^*_{..}-\mu)^2$$

$$= (N+N^*)\gamma_a^{-1}[\mu-(N\bar{y}_{..}+N^*\bar{y}^*_{..})/(N+N^*)]^2 + \gamma_a^{-1}H$$

with

$$H = N\bar{y}_{..}^{\;2} + N^*\bar{y}^{*\;2}_{..} - (N\bar{y}_{..}+N^*\bar{y}^*_{..})^2/(N+N^*)$$

$$= NN^*(N+N^*)^{-1}(\bar{y}_{..}+\bar{y}^*_{..})^2 \qquad \text{(III.E.12)}$$

and $N = IJ$.

The posterior distribution (III.E.11) can be rewritten as

$$p_{11}(\mu,\gamma_a,\gamma_e|\underset{\sim}{y}) = p_{12}(\mu|\gamma_a,\gamma_e,\underset{\sim}{y})p_{13}(\gamma_a,\gamma_e|\underset{\sim}{y}), \qquad \text{(III.E.13)}$$

where

$$p_{12}(\mu|\gamma_a,\gamma_e,\underset{\sim}{y}) \propto \gamma_a^{-1/2}$$

$$\cdot \exp\{-(1/2)(N+N^*)\gamma_a^{-1}[\mu-\frac{(N\bar{y}_{..}+N^*\bar{y}^*_{..})}{(N+N^*)}]^2\}$$

$$(-\infty<\mu<\infty) \qquad \text{(III.E.14)}$$

and

$$p_{13}(\gamma_a,\gamma_e|\underset{\sim}{y}) \propto (\gamma_a)^{-[(1/2)(f_a+f_a^*+1)+1]}$$

$$\cdot (\gamma_e)^{-[(1/2)(f_e+f_e^*)+1]}$$

$$\cdot \exp\{-(1/2)[(SS_e+SS_e^*)/\gamma_e + (SS_a+SS_a^*+H)/\gamma_a]\}$$

$$(\gamma_a \geq \gamma_e > 0). \qquad \text{(III.E.15)}$$

Note that the marginal posterior p.d.f. of γ_a and γ_e is given by (III.E.15) and has the same form as (III.E.3). The quantity f_a^* in (III.E.3) becomes f_a^*+1 in (III.E.15) and SS_a^* becomes $SS_a^* + H$. Thus, from (III.E.15), we obtain the following estimators analogous to estimators E13-E16:

$$\text{E17:}\quad \hat{\rho} = \frac{(SS_e+SS_e^*)(I+f_a^*-2)}{(SS_a+SS_a^*+H)(IJ-I-f_e^*+2)}$$

$$(I+f_a^*>2) \qquad \text{(III.E.16)}$$

(mode of untruncated marginal posterior distribution),

$$\text{E18:}\ \hat{\rho} = \min\{\text{E17},\ 1\} \qquad (I+f_a^*>2) \qquad \text{(III.E.17)}$$

(mode of truncated marginal posterior distribution),

$$\text{E19:}\ \hat{\rho} = \frac{(SS_e+SS_e^*)(I+f_a^*)}{(SS_a+SS_a^*+H)(IJ-I+f_e^*-2)}$$

$$(IJ-I+f_e^*>2) \qquad \text{(III.E.18)}$$

(mean of untruncated marginal posterior distribution), and

$$\text{E20:}\quad \hat{\rho} = \frac{(1-x)}{x}\,\frac{c}{(d-1)}\,\frac{I_x(c+1,d-1)}{I_x(c,d)}$$

$$(IJ-I+f_e^*>2) \qquad \text{(III.E.19)}$$

(mean of truncated marginal posterior distribution). Here,

$$x = (SS_a+SS_a^*+H)/(SS_e+SS_e^*+SS_a+SS_a^*+H)$$

$$c = (1/2)(f_a+f_a^*+1) = (1/2)(I+f_a^*) \qquad . \qquad \text{(III.E.20)}$$

$$d = (1/2)(f_e+f_e^*) = (1/2)(IJ-I+f_e^*)$$

From (III.E.14), we find that the posterior distribution of μ, conditional on γ_a and γ_e, is

$$\mu \sim N[(N\bar{y}_{..}+N^*\bar{y}^*_{..})/(N+N^*),\quad \gamma_a/(N+N^*)]. \tag{III.E.21}$$

Thus, the conditional posterior mean of μ, given γ_a and γ_e, is

$$E(\mu|\gamma_a,\gamma_e,\underline{y}) = (N\bar{y}_{..}+N^*\bar{y}^*_{..})/(N+N^*). \tag{III.E.22}$$

Since this expression does not depend on γ_a and γ_e, it equals the unconditional posterior mean. Thus, the Bayes estimator of μ is

$$\hat{\mu} = (N\bar{y}_{..}+N^*\bar{y}^*_{..})/(N+N^*). \tag{III.E.23}$$

Note that the Bayes estimators of μ_i, α_i, and μ are not of the forms (II.B.20), (II.B.21), and (II.B.22). Our results suggest that consideration be given to estimators of the form

$$\hat{\alpha}_i = (1-\hat{\rho})(\bar{y}_{i.}-\hat{\mu}),\quad \hat{\mu}_i = \hat{\alpha}_i+\hat{\mu}, \tag{III.E.24}$$

where $\hat{\mu}$ is given by expression (III.E.23) and $\hat{\rho}$ is one of the estimators E17-E20.

There are many possible choices for the prior distribution of μ, γ_a, and γ_e in addition to those considered in this section and in section III.D. Naqvi (1969, Chapter III), Zacks (1967), Hill (1977), and Klotz et al. (1969) discuss various alternative prior distributions for μ, γ_a, and γ_e.

F. Summary and Classification of Estimators

A complete list of the estimators of ρ considered in this chapter is as follows:

E1: $\hat{\rho} = \frac{SS_e}{SS_a}\frac{(I-1)}{(IJ-I)}$ $(I>1,\ J>1)$.

E2: $\hat{\rho} = \min\{E1,\ 1\}$.

E3: $\hat{\rho} = \frac{SS_e}{SS_a}\frac{1}{(J-1)}$ $(J>1)$.

E4: $\hat{\rho} = E1$.

E5: $\hat{\rho} = \frac{SS_e}{SS_a}\frac{(I+1)}{(IJ-I+2)}$

E6: $\hat{\rho} = \min\{E3,\ 1\}$.

E7: $\hat{\rho} = E2$.

E8: $\hat{\rho} = \min\{E5,\ 1\}$.

E9: $\hat{\rho} = \frac{SS_e}{SS_a}\frac{(I-3)}{(IJ-I+2)}$ $(I>3)$.

E10: $\hat{\rho} = \frac{SS_e}{SS_a}\frac{(I-1)}{(IJ-I-2)}$ $(I>1,\ IJ-I>2)$.

E11: $\hat{\rho} = \min\{E9,\ 1\}$.

E12: $\hat{\rho} = \frac{c}{(d-1)}\frac{(1-x)}{x}\frac{I_x(c+1,d-1)}{I_x(c,d)}$ $(I>1,\ IJ-I>2)$,

with $x = SS_a/(SS_e+SS_a)$, $c = (1/2)(I-1)$, and $d = (1/2)I(J-1)$.

E13: $\hat{\rho} = \dfrac{(SS_e+SS_e^*)(I+f_a^*-3)}{(SS_a+SS_a^*)(IJ-I+f_e^*+2)}$ $\quad (I+f_a^*>3)$.

E14: $\hat{\rho} = \min\{E13, 1\}$.

E15: $\hat{\rho} = \dfrac{(SS_e+SS_e^*)(I+f_a^*-1)}{(SS_a+SS_a^*)(IJ-I+f_e^*-2)}$ $\quad (I+f_a^*>1,\ IJ-I+f_e^*>2)$.

E16: $\hat{\rho} = \dfrac{c}{(d-1)}\,\dfrac{(1-x)}{x}\,\dfrac{I_x(c+1,d-1)}{I_x(c,d)}$ $\quad (c>0,\ d>1)$,

with $x = (SS_a+SS_a^*)/(SS_e+SS_e^*+SS_a+SS_a^*)$, $c = (1/2)(I-1+f_a^*)$, and

$d = (1/2)(IJ-I+f_e^*)$.

E17: $\hat{\rho} = \dfrac{(SS_e+SS_e^*)}{(SS_a+SS_a^*+H)}\,\dfrac{(I+f_a^*-2)}{(IJ-I+f_e^*+2)}$ $\quad (I+f_a^*>2)$,

with $H = NN^*(N+N^*)^{-1}(\bar{y}_{..}+\bar{y}^*_{..})^2$, $N = IJ$, and $N^* = f_a^*+f_e^*+1$.

E18: $\hat{\rho} = \min\{E17, 1\}$.

E19: $\hat{\rho} = \dfrac{(SS_e+SS_e^*)(I+f_a^*)}{(SS_a+SS_a^*+H)(IJ-I+f_e^*-2)}$ $\quad (IJ-I+f_e^*>2)$,

with H as defined in E17.

E20: $\hat{\rho} = \dfrac{c}{(d-1)}\,\dfrac{(1-x)}{x}\,\dfrac{I_x(c+1,d-1)}{I_x(c,d)}$ $\quad (c>0,\ d>1)$,

with $x = (SS_a+SS_a^*+H)/(SS_e+SS_e^*+SS_a+SS_a^*+H)$, $c = (1/2)(I+f_a^*)$,

$d = (1/2)(IJ-I+f_e^*)$, and with H as defined in E17.

There are two repetitions on this list (E4=E1 and E7=E2), so the total number of different estimators is 18.

The corresponding estimators of μ, α_i, μ_i, and $\tau = \ell_0\mu + \sum_h \ell_h\alpha_h$ are given by

$$\hat{\mu} = \hat{\hat{\mu}} = \bar{y}_{..}$$

$$\hat{\alpha}_i = (1-\hat{\rho})(\bar{y}_{i.}-\bar{y}_{..}) \quad (i = 1,\ldots,I) \qquad \text{(III.F.1)}$$

$$\hat{\mu}_i = \hat{\alpha}_i + \bar{y}_{..} \quad (i = 1,\ldots,I)$$

$$\hat{\tau} = \ell_0\bar{y}_{..} + \sum_h \ell_h\hat{\alpha}_h$$

for estimators E1-E16, and by

$$\hat{\mu} = \frac{(N\bar{y}_{..}+N^*\bar{y}^*_{..})}{(N+N^*)} \quad (N=IJ,\ N^*=f_e^*+f_a^*+1)$$

$$\hat{\alpha}_i = (1-\hat{\rho})(\bar{y}_{i.}-\hat{\mu}) \quad (i=1,\ldots,I) \qquad \text{(III.F.2)}$$

$$\hat{\mu}_i = \hat{\alpha}_i + \bar{y}_{..} \quad (i=1,\ldots,I)$$

$$\hat{\tau} = \ell_0\hat{\mu} + \sum_h \ell_h\hat{\alpha}_h$$

for estimators E17-E20.

To facilitate our investigation of the properties of the estimators of μ, α_i, μ_i τ, and $\underline{\mu}$ it is helpful to classify them according to "type". Note that all of the estimators of ρ depend on the data only through the set of complete sufficient statistics SS_a, SS_e, and $\bar{y}_{..}$. We define five types (classes) of estimators of μ, α_i, μ_i, τ, and $\underline{\mu}$ as follows:

Type A: This class consists of estimators of the form

$$\hat{\mu}_A = \hat{\mu} = \bar{y}_{..},$$

$$\hat{\alpha}_{A;i} = \hat{\alpha}_{A,k;i} = (1-\hat{\rho}_{A,k})(\bar{y}_{i.}-\bar{y}_{..}) \qquad (i = 1,\ldots,I),$$

$$\hat{\mu}_{A;i} = \hat{\mu}_{A,k;i} = \bar{y}_{..} + \hat{\alpha}_{A,k;i} \qquad (i = 1,\ldots,I),$$

and

$$\hat{\tau}_A = \hat{\tau}_{A,k} = \ell_0 + \bar{y}\sum_h \ell_h \hat{\alpha}_{A,k;h},$$

with

$$\hat{\rho}_A = \hat{\rho}_{A,k} = k\ SS_e/SS_a,$$

where k is an arbitrary positive constant. Let $\underline{\hat{\mu}}_A = \underline{\hat{\mu}}_{A,k}$ denote the vector of dimensions Ix1 whose ith component is $\hat{\mu}_{A,k;i}$.

Type A estimator will be called untruncated estimators. Estimators E1, E3, E4, E5, E7, and E9 are in this class.

Type B: This class consists of estimators of the form

$$\hat{\mu}_B = \hat{\mu} = \bar{y}_{..},$$

$$\hat{\alpha}_{B;i} = \hat{\alpha}_{B,k;i} = (1-\hat{\rho}_{B,k})(\bar{y}_{i.}-\bar{y}_{..}) \qquad (i = 1,\ldots,I),$$

$$\hat{\mu}_{B;i} = \hat{\mu}_{B,k;i} = \bar{y}_{..} + \hat{\alpha}_{B,k;i} \qquad (i = 1,\ldots,I),$$

and

$$\hat{\tau}_B = \hat{\tau}_{B,k} = \ell_0\bar{y}_{..} + \sum_h \ell_h \hat{\alpha}_{B,k;h},$$

with

$$\hat{\rho}_B = \hat{\rho}_{B,k} = \min\{k\ SS_e/SS_a,\ 1\},$$

where k is an arbitrary positive constant. Let $\underline{\hat{\mu}}_B = \underline{\hat{\mu}}_{B,k}$ denote the vector of dimensions Ix1 whose ith element is $\hat{\mu}_{B,k;i}$.

Type B estimators will be called <u>truncated</u> estimators. Estimators E2, E6, E7, E8, and E11 are in this class.

<u>Type C</u>: This class consists of estimators of the form

$$\hat{\mu}_C = \hat{\mu} = \bar{y}_{..},$$

$$\hat{\alpha}_{C,i} = (1-\hat{\rho}_C)(\bar{y}_{i.}-\bar{y}_{..}) \qquad (i = 1,\ldots,I),$$

$$\hat{\mu}_{C;i} = \bar{y}_{..} + \hat{\alpha}_{C;i} \qquad (i = 1,\ldots,I),$$

$$\hat{\tau}_C = \ell_0\bar{y}_{..} + \sum_h \ell_h\hat{\alpha}_{C;h},$$

with

$$\hat{\rho}_C = f_C(SS_e/SS_a),$$

where $f_C(x)$ is an arbitrary positive function of $x>0$. Let $\hat{\mu}_C$ denote the vector of dimensions Ix1 whose ith element is $\hat{\mu}_{C;i}$.

Note that estimators of Types A and B are special cases of Type C estimators. In addition to estimators of Types A and B, this class includes estimator E12.

<u>Type D</u>: This class consists of estimators of the form

$$\hat{\mu}_D = \hat{\mu} = \bar{y}_{..},$$

$$\hat{\alpha}_{D;i} = (1-\hat{\rho}_D)(\bar{y}_{i.}-\bar{y}_{..}) \qquad (i = 1,\ldots,I),$$

$$\hat{\mu}_{D;i} = \bar{y}_{..} + \hat{\alpha}_{D;i} \qquad (i = 1,\ldots,I),$$

and

$$\hat{\tau}_D = \ell_0\bar{y}_{..} + \sum_h \ell_h\hat{\alpha}_{D;h},$$

with

$$\hat{\rho}_D = f_D(SS_e, SS_a),$$

where $f_D(x_1,x_2)$ is an arbitrary positive function of $x_1>0$ and $x_2>0$. Let $\underline{\hat{\mu}}_D$ denote the vector of dimensions Ix1 whose ith element is $\hat{\mu}_{D;i}$.

Note that Type C estimators are special cases of Type D estimators. In addition to estimators of Type C, this class includes estimators E13, E14, E15, and E16.

<u>Type E</u>: This class consists of estimators of the form

$$\hat{\mu}_E = \frac{(N\bar{y}_{..}+N^*\bar{y}^*_{..})}{(N+N^*)},$$

$$\hat{\alpha}_{E;i} = (1-\hat{\rho}_E)(\bar{y}_{..}-\hat{\mu}_E) \qquad (i = 1,\ldots,I),$$

$$\hat{\mu}_{E;i} = \hat{\mu}_E + \hat{\alpha}_{E;i}$$

and

$$\hat{\tau}_E = \ell_0\hat{\mu}_E + \sum_h \ell_h\hat{\alpha}_{E;h},$$

where $\hat{\rho}_E$ depends (nontrivially) on H as well as SS_e and SS_a, with $H = NN^*(N+N^*)^{-1}(\bar{y}_{..}+\bar{y}^*_{..})^2$, N=IJ, and

$N^* = f_a^* + f_e^* + 1$. Let $\hat{\underline{\mu}}_E$ denote the vector of dimensions $I \times 1$ whose ith element is $\hat{\mu}_{E;i}$.

Estimators E17, E18, E19, and E20 are in this class.

IV. UNCONDITIONAL PROPERTIES OF ESTIMATORS OF RANDOM EFFECTS

A. General Results

Take the model to be the balanced one-way random model. If ρ is known, then the best linear unbiased estimator of α_i is

$$\hat{\hat{\alpha}}_i = (1-\rho)(\bar{y}_{i.}-\bar{y}_{..}) \quad (i = 1,\ldots,I),$$

as discussed in Chapter II. If ρ is unknown, then we can consider any of the estimators of α_i listed in section III.F.

While, in general, Type D estimators of α_i are not linear, they are unbiased under certain regularity conditions, as indicated in the following theorem.

Theorem IV.A.1: Take the model to be the balanced one-way random model. Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ and let $\hat{\alpha}_{D;i}$ be the corresponding Type D estimator of α_i $(i = 1,\ldots,I)$. Then, if $E(\hat{\alpha}_{D;1})$ exists, $\hat{\alpha}_{D;i}$ estimates α_i unbiasedly.

Proof: We need to prove that $E(\hat{\alpha}_{D;i}) = E(a_i)$. By assumption, $E(a_i) = 0$. Note that $\hat{\alpha}_{D;i}$ has the same distribution for all i $(i = 1,\ldots,I)$, and hence, the same expectation. In consequence, using the assumption that $E(\hat{\alpha}_{D;1})$ exists,

$$E(\hat{\alpha}_{D;i}) = E(\hat{\alpha}_{D;1}) = E[(1-\hat{\rho}_D)(\bar{y}_{1.}-\bar{y}_{..})]$$

$$= (1/I)\sum_i E[(1-\hat{\rho}_D)(\bar{y}_{i.}-\bar{y}_{..})]$$

$$= (1/I)E[(1-\hat{\rho}_D)\sum_i(\bar{y}_{i.}-\bar{y}_{..})] = 0,$$

since

$$\sum_i(\bar{y}_{i.}-\bar{y}_{..}) = 0.$$

Q.E.D.

The type of "symmetry argument" used to prove Theorem IV.A.1 will be used repeatedly in our development.

Let $\tau = \ell_0\mu + \sum_i \ell_i\alpha_i$ represent an arbitrary linear combination of μ, $a_1,\ldots,a_I$, and let t be as given by (I.A.16). The statistic $\bar{y}_{..}$ is obviously unbiased for μ; thus, the following corollary is easily obtained.

<u>Corollary IV.A.1</u>: Take the model to be the balanced one-way random model. Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\hat{\tau}_D$ and $\hat{\alpha}_{D;1}$ be the corresponding Type D estimators of τ and α_1. Then, if $E(\hat{\alpha}_{D;1})$ exists, $\hat{\tau}_D$ estimates τ unbiasedly.

Corollary IV.A.1 implies that estimators of τ of Types A, B, C, and D are unbiased, provided that their expectations exist. Type E estimators of τ are obviously biased.

The following theorem is very useful in finding mean squared errors of estimators of linear combinations of μ and the α_i's.

Theorem IV.A.2: Take the model to be the one-way random model. Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\hat{\tau}_D$ and $\hat{\alpha}_{D;1}$ be the corresponding Type D estimators of τ and α_1. Then,

$$\begin{aligned}
\mathrm{MSE}(\hat{\tau}_D,\tau) &= E(\hat{\tau}_D - t)^2 \\
&= \ell_0^2\sigma_e^2(\rho IJ)^{-1} + \mathrm{MSE}(\hat{\alpha}_{D;1},\alpha_1)\sum_i \ell_i^2 \\
&\quad - 2\ell_0\sigma_e^2(1-\rho)(\rho IJ)^{-1}\sum_i \ell_i \\
&\quad + 2(I-1)^{-1}[\sigma_a^2 - \mathrm{MSE}(\hat{\alpha}_{D;1},\alpha_1)]\sum_i \sum_{i'>i} \ell_i\ell_{i'}.
\end{aligned}$$

We introduce two lemmas that can be used to prove Theorem IV.A.2.

Lemma IV.A.1: Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\hat{\alpha}_{D;i}$ be the corresponding Type D estimator of α_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$\begin{aligned}
E[(\hat{\alpha}_{D;i}-a_i)(\hat{\alpha}_{D;i'}-a_{i'})] &= (I-1)^{-1}[\sigma_a^2 - \mathrm{MSE}(\hat{\alpha}_{D;1},\alpha_1)] \\
&= (I-1)^{-1}[\sigma_e^2(1-\rho)(\rho IJ)^{-1} - \mathrm{MSE}(\hat{\alpha}_{D;1},\alpha_1)] \\
&\qquad (i>i' = 1,\ldots,I).
\end{aligned}$$

<u>Proof</u>: Let $\hat{\beta} = \hat{\beta}_D$ and $\hat{\alpha}_i = \hat{\alpha}_{D;i}$. By symmetry, we have that

$$E[(\hat{\alpha}_i - a_i)(\hat{\alpha}_{i'} - a_{i'})] = (I-1)^{-1} \sum_{\substack{i=1 \\ i \neq i'}}^{I} E[(\hat{\alpha}_i - a_i)(\hat{\alpha}_{i'} - a_{i'})]$$

$$= (I-1)^{-1} \sum_{i=1}^{I} E[(\hat{\alpha}_i - a_i)(\hat{\alpha}_{i'} - a_{i'})] - (I-1)^{-1} E(\hat{\alpha}_{i'} - a_{i'})^2$$

$$= (I-1)^{-1} E\{(\hat{\alpha}_1 - a_1) \sum_i [(1-\hat{\beta})(\bar{y}_{i.} - \bar{y}_{..}) - a_i]\}$$

$$- (I-1)^{-1} \mathrm{MSE}(\hat{\alpha}_1, \alpha_1)$$

$$= (I-1)^{-1} \{-E[(\hat{\alpha}_1 - a_1) \sum_i a_i] - \mathrm{MSE}(\hat{\alpha}_1, \alpha_1)\}$$

$$= (I-1)^{1} \{E[(1-\hat{\beta})(\bar{y}_{1.} - \bar{y}_{..}) - a_1][-\sum_i a_i] - \mathrm{MSE}(\hat{\alpha}_1, \alpha_1)\}$$

$$= (I-1)^{-1} \{E[(1-\hat{\beta})(\bar{y}_{1.} - \bar{y}_{..})(-\sum_i a_i)]$$

$$+ \sigma_a^2 - \mathrm{MSE}(\hat{\alpha}_1, \alpha_1)\}$$

But, by symmetry,

$$E[(1-\hat{\beta})(\bar{y}_{1.} - \bar{y}_{..})(-\sum_i a_i)]$$

$$= (1/I) \sum_{i'} E[(1-\hat{\beta})(\bar{y}_{i'.} - \bar{y}_{..})(-\sum_i a_i)] = 0.$$

Q.E.D.

Lemma IV.A.2: Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\hat{\alpha}_{D;i}$ be the corresponding Type D estimator of α_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$E[(\bar{y}_{..}-\mu)(\hat{\alpha}_{D;i}-a_i)] = -\sigma_a^2/I = -\sigma_e^2(1-\rho)(\rho IJ)^{-1}.$$

Proof: By symmetry, we have that

$$E[(\bar{y}_{..}-\mu)\hat{\alpha}_{D;i}] = (1/I)E[(\bar{y}_{..}-\mu)\sum_i\hat{\alpha}_{D;i}] = 0.$$

Thus,

$$E[(\bar{y}_{..}-\mu)(\hat{\alpha}_{D;i}-a_i)] = -E[(\bar{y}_{..}-\mu)a_i] = -E[(\bar{a}_{.}+\bar{e}_{..})a_i]$$

$$= -\sigma_a^2/I.$$

Q.E.D.

Theorem IV.A.2 is an obvious consequence of (II.B.17) and Lemmas IV.A.1 and IV.A.2.

The following theorem is useful in deriving an expression for $MSE(\hat{\alpha}_{D;1},\alpha_1)$.

Theorem IV.A.3: Let $\hat{\rho}_D$ represent an arbitrary Type D estimator of ρ. Then, under the balanced one-way random model, $[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i]$ is distributed independently of $[(\hat{\rho}_D-\rho)(\bar{y}_{i'}-\bar{y}_{..})]$ $(i,i' = 1,\ldots,I)$, and

$$E\{[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i][(\hat{\rho}_D-\rho)(\bar{y}_{i'.}-\bar{y}_{..})]\}=0$$

$$(i,i' = 1,\ldots,I). \qquad \text{(IV.A.1)}$$

Proof: The quantity $(\hat{\rho}_D-\rho)(\bar{y}_{i'.}-\bar{y}_{..})$ is a function of the error contrasts (III.C.3). Therefore, it suffices to prove that $[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i]$ is distributed independently of the error contrasts. Let

$$d_{i,i'} = 1/(I-1), \text{ if } i=i',$$
$$= -1, \qquad \text{if } i\neq i'.$$

We find that

$$E\{[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i](\bar{y}_{i'.}-\bar{y}_{..})\}$$

$$= E\{[(1-\rho)(a_i-\bar{a}_{.}+\bar{e}_{i.}-\bar{e}_{..})-a_i][a_{i'}-\bar{a}_{.}+e_{i'j}-\bar{e}_{..}]$$

$$= d_{i,i'}\{(1-\rho)[(I)^{-1}\sigma_a^2 + (IJ)^{-1}\sigma_e^2] - (I)^{-1}\sigma_a^2\}$$

$$= d_{i,i'}(IJ)^{-1}\{(1-\rho)\sigma_e^2(\rho)^{-1}-\sigma_e^2(\rho)^{-1}(1-\rho)\} = 0$$

for all $i,i' = 1,\ldots,I$ and $j = 1,\ldots,J$. Thus, since $[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i]$ and $(y_{i'j}-\bar{y}_{..})$ are normally distributed and are uncorrelated, they are distributed independently.

Result (IV.A.1) follows from the independence result.

Q.E.D.

<u>Theorem IV.A.4</u>: Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\hat{\alpha}_{D;i}$ be the corresponding Type D estimator of α_i. Then, under the balanced one-way random model,

$$\begin{aligned} MSE(\hat{\alpha}_{D;i},\alpha_i) &= \sigma_e^2\, MSE(\hat{\alpha}_i,\alpha_i) + (IJ)^{-1}E[(\hat{\rho}_D-\rho)^2 SS_a] \\ &= \sigma_e^2(IJ\rho)^{-1}[1+\rho(I-1)] \\ &\quad + (IJ)^{-1}E[(\hat{\rho}_D-\rho)^2 SS_a] \\ &\quad (i = 1,\ldots,I). \end{aligned} \tag{IV.A.2}$$

<u>Proof</u>: Using Theorem IV.A.3, we find that

$$\begin{aligned} MSE(\hat{\alpha}_{D;i},\alpha_i) &= E(\hat{\alpha}_{D,i}-a_i)^2 \\ &= E[(1-\hat{\rho}_D)(\bar{y}_{i.}-\bar{y}_{..})-a_i]^2 \\ &= E[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i-(\hat{\rho}_D-\rho)(\bar{y}_{i.}-\bar{y}_{..})]^2 \\ &= E[(1-\rho)(\bar{y}_{i.}-\bar{y}_{..})-a_i]^2 + E[(\hat{\rho}_D-\rho)^2(\bar{y}_{i.}-\bar{y}_{..})^2]. \end{aligned} \tag{IV.A.3}$$

The first term of (IV.A.3) is the mean squared error of $\hat{\alpha}_i$, the BLUE of α_i when ρ is known, and hence is equal to expression (II.B.16). By symmetry, the second term of (IV.A.3) is equal to

$$(IJ)^{-1}\sum_i\sum_j E[(\hat{\rho}_D-\rho)^2(\bar{y}_{i.}-\bar{y}_{..})^2] = (IJ)^{-1}E[(\hat{\rho}_D-\rho)^2 SS_a].$$

Q.E.D.

The second term of (IV.A.2) can be interpreted as a "penalty" that we pay for not knowing the value of ρ.

Theorem IV.A.5: Let $\hat{\rho}_C = f_C(SS_e/SS_a)$ where $f_C(x)$ is an arbitrary positive function of $x>0$. Then, under the balanced one-way random model,

$$E[(\hat{\rho}_C-\rho)^2 SS_a] = \sigma_e^2(1/\rho)(IJ-1)E\{[f_C(\tfrac{\rho s}{1-s})-\rho]^2(1-s)\},$$

where s is a beta random variable with parameters $I(J-1)/2$ and $(I-1)/2$.

Proof: Let $\nu_e = I(J-1)/2$, $\nu_a = (I-1)/2$, $u_e = SS_e/\gamma_e$, $u_a = SS_a/\gamma_a$, $w = u_e+u_a$, and $s = u_e/(u_e+u_a)$. The random variables u_e and u_a are distributed independently as chi-square random variables with degrees of freedom $I(J-1)$ and $(I-1)$, respectively. Thus, w has a chi-square distribution with $IJ-1$ degrees of freedom, s has a beta distribution with parameters ν_e and ν_a, and w and s are distributed independently (e.g., Johnson and Kotz, 1970, section 24.2).

Therefore,

$$SS_e/SS_a = (\gamma_e/\gamma_a)(u_e/u_a) = \rho s(1-s)^{-1}, \text{ and}$$

$$SS_a = (1-s)w\gamma_a = (1-s)w\sigma_e^2/\rho.$$

Thus,

$$E[(\hat{\rho}_C-\rho)^2 SS_a] = E\{[f_C(\tfrac{\rho s}{1-s})-\rho]^2(1-s)w\gamma_a\}$$

$$= \gamma_a E(w)E\{[f_C(\tfrac{\rho s}{1-s})-\rho]^2(1-s)\}.$$

Further,

$$\gamma_a E(w) = \sigma_e^2 (1/\rho)(IJ-1).$$

Q.E.D.

Note that Theorem IV.A.5 covers estimators of Types A, B, and C.

Theorem IV.A.6: Take the model to be the balanced one-way random model. Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ and let $\underline{\hat{\mu}}_D$ and $\hat{\mu}_{D;i}$ be the corresponding Type D estimators of $\underline{\mu}$ and μ_i $(i = 1,\ldots,I)$. Then, if

$$E(\hat{\mu}_{D;1}) \text{ exists,}$$
$$TB(\underline{\hat{\mu}}_D,\underline{\mu}) = TB(\underline{\hat{\mu}}_D,\underline{\mu}) = 0.$$

Theorem IV.A.6 is an immediate consequence of Theorem IV.A.1.

Theorem IV.A.7: Let $\hat{\rho}_D$ be an arbitrary Type D estimator of ρ, and let $\underline{\hat{\mu}}_D$, $\hat{\mu}_{D;i}$, and $\hat{\alpha}_{D;i}$ be the corresponding Type D estimators of $\underline{\mu}$, μ_i, and α_i $(i = 1,\ldots,I)$, respectively. Then, under the balanced one-way random model,

(i) $$TMSE(\underline{\hat{\mu}}_D,\underline{\mu}) = J\sum_i E(\hat{\mu}_{D;i}-\hat{\mu}_i)^2 = (IJ)MSE(\hat{\mu}_{D;1},\mu_1),$$

(ii) $$MSE(\hat{\mu}_{D;i},\mu_i) = MSE(\hat{\alpha}_{D;1},\alpha_1) + \frac{\sigma_e^2}{IJ}\,\frac{(2\rho-1)}{\rho}$$

$$(i = 1,\ldots,I).$$

Part (ii) of Theorem IV.A.7 is an immediate consequence of Theorem IV.A.2. Part (i) follows from the symmetry of the distributions of the $\hat{\beta}_{D;i}$'s.

In section IV.B, we derive a simple closed-form expression for mean squared errors of Type A estimators of the α_i's. It is not possible to obtain a closed-form expression for the mean squared errors of estimators of Types B, C, D, and E, though, in section IV.C, we express the mean squared errors of the Type B estimators in terms of the incomplete beta function ratio. Theorems IV.A.4 and IV.A.5 imply that the computation of MSEs of Type C estimators can be reduced to the numerical evaluation of a one-dimensional integral, while the computation of MSEs of Type D or E estimators requires the numerical evaluation of two- or three-dimensional integrals. However, we limit our discussion to the mean squared errors of estimators of Type A or B.

B. Unconditional Properties of Type A Estimators

Corollary IV.A.1 implies that Type D estimators of an arbitrary parametric function τ are unbiased, provided that their expectations exist. We now show that if $I>3$, then Type A (untruncated) estimators of τ satisfy this requirement.

<u>Theorem IV.B.1</u>: Take the model to be the balanced one-way random model. Let $\hat{\rho}_A$ be an arbitrary Type A estimator of ρ, and let $\hat{\alpha}_{A;1}$ be the corresponding estimator of α_1. Then, if $I>3$, $E[|\hat{\alpha}_{A;1}|] < \infty$.

<u>Proof</u>: For any random variable X, we have that

$$E[|X|] \leq [E(X^2)]^{1/2}. \qquad \text{(IV.B.1)}$$

Inequality (IV.B.1) is a special case of Hölder's inequality (e.g., Chung, 1974, page 47). Note that

$$E[|\hat{\alpha}_{A.1}|] = E[|1-\hat{\rho}_A|\cdot|\bar{y}_{1.}-\bar{y}_{..}|] < \infty$$

if and only if $E[\hat{\rho}_A|\bar{y}_{1.}-\bar{y}_{..}|] < \infty$. Thus, using (IV.B.1),

$$E[\hat{\rho}_A|\bar{y}_{1.}-\bar{y}_{..}|] \leq \{E[\rho_A{}^2(\bar{y}_{1.}-\bar{y}_{..})^2]\}^{1/2}$$

$$= \{(IJ)^{-1}\sum_i \sum_j E[\rho_A{}^2(\bar{y}_{i.}-\bar{y}_{..})^2]\}^{1/2}$$

$$= \{(IJ)^{-1}E[\rho_A{}^2 SS_a]\}^{1/2}$$

$$= \{(IJ)^{-1}E[k^2 SS_e{}^2/SS_a]\}^{1/2}$$

$$= \{k^2(\gamma_e{}^2/\gamma_a)I(J-1)(IJ-I+2)/(I-3)\}^{1/2} < \infty, \qquad \text{(IV.B.2)}$$

since $I>3$.

Q.E.D.

Corollary IV.B.1: Take the model to be the balanced one-way random model and let $\hat{\tau}_A$ be an arbitrary Type A estimator of τ. Then, if $I>3$, $\hat{\tau}_A$ estimates τ unbiasedly.

Corollary IV.B.1 is an immediate consequence of Theorem IV.B.1 and Corollary IV.A.1.

By applying the results of section IV.A, it is easy to derive simple expression for MSEs of untruncated estimators of α_i.

Theorem IV.B.2: Take the model to be the balanced one-way random model. Let k be an arbitrary positive constant and let $\hat{\alpha}_{A,k;i}$ be the corresponding Type A estimator of α_i $(i = 1,\ldots,I)$. Then,

$$\begin{aligned}\mathrm{MSE}(\hat{\alpha}_{A,k;i},\alpha_i) &= \mathrm{MSE}(\hat{\alpha}_{A,k;1},\alpha_1)\\ &= \sigma_e^2(1-\rho)(\rho IJ)^{-1}[1+\rho(I-1)]\\ &\quad + \sigma_e^2\rho(IJ)^{-1}\{k^2I(J-1)(IJ-I+2)(I-3)^{-1}\\ &\quad - 2kI(J-1)+I-1\}\end{aligned}$$

$$(I>3;\quad i = 1,\ldots,I). \qquad \text{(IV.B.3)}$$

Proof: Let $\hat{\rho} = \hat{\rho}_{A,k}$ be the corresponding untruncated estimator of ρ. We find that

$$E[(\hat{\rho}-\rho)^2 SS_a] = E[(kSS_e/SS_a-\rho)^2 SS_a]$$

$$= k^2 E(SS_e^2/SS_a) - 2k\rho E(SS_e) + \rho^2 E(SS_a)$$

$$= k^2\gamma_e^2 I(J-1)(IJ-I+2)[\gamma_a(I-3)]^{-1} - 2k\rho\gamma_e I(J-1) + \rho^2\gamma_a(I-1)$$

$$= \sigma_e^2\rho\{k^2 I(J-1)(IJ-I+2)(I-3)^{-1} - 2kI(J-1) + I-1\}$$

Substituting this expression into expression (IV.A.2), we obtain expression (IV.B.3).

Q.E.D.

Theorems IV.A.2 and IV.B.2 can be used to obtain expressions for MSEs of untruncated estimators of an arbitrary linear combination of μ, $a_1,\ldots,a_I$. The following corollary is an important special case of the general result.

Corollary IV.B.2: Take the model to be the balanced one-way random model. Let k be an arbitrary positive constant and let $\hat{\mu}_{A,k;i}$ be the corresponding Type A estimator of μ_i $(i = 1,\ldots,I)$. Then,

$$MSE(\hat{\mu}_{A,k;i},\mu_i) = \frac{\sigma_e^2}{IJ}\{I+k\rho I(J-1)[\frac{k(IJ-I+2)}{(I-3)} - 2]\}$$

$$(I>3; \quad i = 1,\ldots,I). \qquad \text{(IV.B.4)}$$

This corollary is obtained by substituting expression (IV.B.3) into the result of Theorem IV.A.7 (ii).

It is easy to show that expression (IV.B.3) is minimized uniformly with respect to k by

$$k^* = (I-3)(IJ-I+2)^{-1}, \qquad \text{(IV.B.5)}$$

which is the value obtained by taking $\hat{\rho}_A$ to be the mode of the untruncated posterior marginal distribution of ρ corresponding to Jeffreys' prior (estimator E9).

For this value of k, (IV.B.3) reduces to:

$$\begin{aligned} MSE(\hat{\alpha}_{A,k^*;i},\alpha_i) = {} & \sigma_e^2(1-\rho)(\rho IJ)^{-1}[1+\rho(I-1)] \\ & + 2\sigma_e^2\rho(IJ)^{-1}(IJ-1)(IJ-I+2)^{-1} \\ & (I>3;\quad i = 1,\ldots,I). \qquad \text{(IV.B.6)} \end{aligned}$$

The first term of this expression [and of expression (IV.B.2)] equals to the mean squared error of $\hat{\alpha}_i$, the quantity that would be the BLUE of α_i if ρ were known. The second term can be interpreted as a penalty that is paid for not knowing the value of ρ.

It follows from Theorem IV.A.7 (ii) that the value of k given by expression (IV.B.5) is also the value that minimizes $MSE(\hat{\mu}_{A,k;i},\mu_i)$. Thus, the optimal Type A estimator of $\mu+\alpha_i$ is

$$\hat{\mu}_{JS;i} = \hat{\mu}_{A,k^*;i} = \bar{y}_{i.} - \frac{(I-3)}{(IJ-I+2)}\frac{SS_e}{SS_a}(\bar{y}_{i.}-\bar{y}_{..})$$
$$(i = 1,\ldots,I). \qquad \text{(IV.B.7)}$$

The vector $\hat{\underline{\mu}}_{JS} = \hat{\underline{\mu}}_{A,k}$, with elements given by expression (IV.B.7), can also be derived by applying the approach of James and Stein (1961) to the problem of estimating the mean vector of the balanced one-way fixed-effects model. It will be shown in Chapter V that the vector $\hat{\underline{\mu}}_{JS}$ is also the optimal, among Type A estimators of $\underline{\mu}$, with respect to the TCMSE criterion.

C. Unconditional Properties of Type B Estimators

As indicated by the following theorem, Type B estimators of an arbitrary linear parametric function τ are unbiased under the same conditions required for unbiasedness of Type A estimators.

Theorem IV.C.1: Take the model to be the balanced one-way random model and let $\hat{\tau}_B$ be an arbitrary Type B estimator of τ. Then, if $I>3$, $\hat{\tau}_B$ estimates τ unbiasedly.

Proof: Let $\hat{\alpha}_{B;1}$ and $\hat{\rho}_B$ be the corresponding Type B estimators of α, and ρ, respectively. Let $\hat{\rho}_A$ be the corresponding estimator of ρ, i.e., $\hat{\rho}_B = \min\{\hat{\rho}_A, 1\}$. According to Corollary IV.A.1, it suffices to prove that $E[|\hat{\alpha}_{B;1}|]<\infty$. Note that $E[|\hat{\alpha}_{B;1}|]<\infty$ if and only if $E[\hat{\rho}_B|\bar{y}_{1.}-\bar{y}_{..}|] < \infty$. Using (IV.B.2), we have that

$$E[\hat{\rho}_B|\bar{y}_{1.}-\bar{y}_{..}|] \le E[\hat{\rho}_A|\bar{y}_{1.}-\bar{y}_{..}|]<\infty .$$

Q.E.D.

Let

$$\nu_e = I(J-1)/2,\ \nu_a = (I-1)/2,\ \text{and } x = (1+k\rho)^{-1}, \qquad \text{(IV.C.1)}$$

where k is an arbitrary positive constant. Let s represent a beta random variable with parameters ν_e and ν_a. Recall that the probability density function of the beta distribution with parameters c and d is

$$h(t;a,b) = B^{-1}(c,d)t^{c-1}(1-t)^{d-1}$$

$$(0<t<1,\quad c>0,\quad d>0). \qquad \text{(IV.C.2)}$$

We have that

$$E\{[\min(\tfrac{\rho ks}{1-s}\ ,\ 1)-\rho]^2(1-s)\}$$

$$= \rho^2E(1-s)-2k\rho^2 \int_0^x th(t;\nu_e,\nu_a)dt$$

$$+ \rho^2k^2 \int_0^x t^2(1-t)^{-1}h(t;\nu_e,\nu_a)dt$$

$$+ (1-2\rho)\int_x^1 (1-t)h(t;\nu_e,\nu_a)dt$$

$$= (\rho^2+1-2\rho)E(1-s)$$

$$- 2k\rho^2 \int_0^x th(t;\nu_e,\nu_a)dt$$

$$+ \rho^2k^2 \int_0^x t^2(1-t)^{-1}h(t;\nu_e,\nu_a)dt$$

$$- (1-2\rho) \int_0^x (1-t)h(t;\nu_e,\nu_a)dt$$

$$= (1-\rho)^2 \frac{\nu_a}{(\nu_e+\nu_a)} - 2k\rho^2 \frac{\nu_e}{(\nu_e+\nu_a)} I_x(\nu_e+1,\nu_a)$$

$$+ \rho^2 k^2 \frac{\nu_e(\nu_e+1)}{(\nu_a-1)(\nu_e+\nu_a)} I_x(\nu_e+2, \nu_a-1)$$

$$- (1-2\rho)\frac{\nu_a}{(\nu_e+\nu_a)} I_x(\nu_e,\nu_a+1) \qquad \text{(IV.C.3)}$$

Making use of the well-known recursive formula

$$I_x(c+1,d-1) = I_x(c,d) - \frac{\Gamma(c+d)}{\Gamma(c+1)\ (d)} x^c(1-x)^{d-1} \qquad \text{(IV.C.4)}$$

(e.g., Abramowitz and Stegun, 1970, section 26.5), we find that

$$E\{[\min(\frac{\rho ks}{1-s}, 1)-\rho]^2(1-s)\} = \Psi, \qquad \text{(IV.C.5)}$$

where

$$\Psi = (1-\rho)^2 \frac{(I-1)}{(IJ-1)} + \frac{I_x(\nu_e,\nu_a+1)}{(IJ-1)}[-2k\rho^2 I(J-1)-(1-2\rho)(I-1)$$

$$+ \rho^2 k^2 \frac{I(J-1)(IJ-I+2)}{(I-3)}]$$

$$+ \frac{x^{\nu_e+1}(1-x)^{\nu_a+1}}{B(\nu_e,\nu_a)}[\frac{\rho}{(I-1)} \frac{-2\rho k(IJ-I+2)}{(I-3)(I-1)} \frac{-2}{(I-3)}]. \qquad \text{(IV.C.6)}$$

Using Theorems IV.A.4 and IV.A.5, and result (IV.C.5), we obtain the following theorem:

<u>Theorem IV.C.2</u>: Take the model to be the balanced one-way random model. Let k be an arbitrary positive constant and let $\hat{\alpha}_{B,k;i}$ be the corresponding Type B estimator of

α_i $(i = 1,\ldots,I)$. Then,

$$\mathrm{MSE}(\hat{\alpha}_{B,k;i},\alpha_i) = \sigma_e^2(IJ\rho)^{-1}\,[1+\rho(I-1)] + (IJ-1)\Psi$$

$$(I>3;\quad i = 1,\ldots,I), \qquad \text{(IV.C.7)}$$

where Ψ, x, ν_e, and ν_a are defined by expressions (IV.C.6) and (IV.C.1).

Note that Ψ and hence expression (IV.C.7) involve the incomplete beta function ratio. Thus, in general, numerical methods will have to be used to evaluate expression (IV.C.7). It can be shown that there is no value of k that minimizes expression (IV.C.7) uniformly.

As indicated by the following theorem, the truncated (Type B) estimator of α_i always dominates the corresponding untruncated (Type A) estimator.

Theorem IV.C.3: Take the model to be the balanced one-way random model. Let $\hat{\rho}_U$ be an arbitrary Type D estimator of ρ. Take $\hat{\rho}_T = \min\{\hat{\rho}_U, 1\}$, and define

$$\hat{\alpha}_{U;i} = (1-\hat{\rho}_U)(\bar{y}_{i.}-\bar{y}_{..}) \text{ and } \hat{\alpha}_{T;i} = (1-\hat{\rho}_T)(\bar{y}_{i.}-\bar{y}_{..})$$

$$(i = 1,\ldots,I).$$

Then,

$$E[(\hat{\alpha}_{T;i}-a_i)^2] \leq E[(\hat{\alpha}_{U;i}-a_i)^2], \qquad \text{(IV.C.8)}$$

with strict inequality if $P(\hat{\rho}_U>1) > 0$ and $E[(\hat{\alpha}_{T;i}-a_i)^2]<\infty$.

<u>Proof</u>: If $E[(\hat{\alpha}_{U;i}-a_i)^2] = \infty$ then inequality (IV.C.8) is obviously satisfied. If $E[(\hat{\alpha}_{U;i}-a_i)^2]<\infty$ then, using Theorem IV.A.4,

$$(IJ)\{E[(\hat{\alpha}_{T;i}-a_i)^2] - E[(\hat{\alpha}_{U;i}-a_i)^2]\}$$

$$= E[(\hat{\rho}_T-\rho)^2 SS_a] - E[(\hat{\rho}_U-\rho)^2 SS_a]$$

$$= E\{[\hat{\rho}_T{}^2-\hat{\rho}_U{}^2-2\rho(\hat{\rho}_T-\hat{\rho}_U)]SS_a\}$$

$$= E\{[1-\hat{\rho}_U{}^2-2\rho(1-\hat{\rho}_U)]SS_a|\hat{\rho}_U > 1\}P(\hat{\rho}_U>1)$$

$$\leq E\{[1-\hat{\rho}_U{}^2-2(1-\hat{\rho}_U)2SS_a|\hat{\rho}_U>1\}P(\hat{\rho}_U>1)$$

$$= -E\{(1-\hat{\rho}_U)^2 SS_a|\hat{\rho}_U>1\}P(\hat{\rho}_U>1) \leq 0.$$

Further, the last inequality is strict if $P(\hat{\rho}_U>1)>0$.
Q.E.D.

As a special case, we find that, for $I>3$, the Type B estimator

$$\hat{\alpha}_{TJS;i} = (1-\hat{\rho}_{TJS})(\bar{y}_{i.}-\bar{y}_{..}) \qquad \text{(IV.C.9)}$$

where

$$\hat{\rho}_{TJS} = \min\{(I-3)(IJ-I+2)^{-1}(SS_e/SS_a),1\}, \qquad \text{(IV.C.10)}$$

has uniformly smaller mean squared error than any Type A estimator of α_i ($i = 1,\ldots,I$). Note that $\hat{\rho}_{TJS}$ is the mode of the truncated marginal posterior distribution of ρ corresponding to Jeffreys' prior distribution (estimator E11).

V. CONDITIONAL PROPERTIES OF ESTIMATORS OF RANDOM EFFECTS

A. Preliminaries

In the present chapter, we describe various conditional properties (i.e., properties conditional on $\underline{a} = \underline{\alpha}$) of the estimators of the group mean $\mu_i = \mu+\alpha_i$ $(i = 1,\ldots,I)$ derived in Chapter III. The model is the balanced one-way random model (I.A.1) and we use the notation introduced in section I.A. The symbol CE denotes conditional expectations, i.e., we define the operator $CE(\cdot)$ to be the same as the operator $E(\cdot|\underline{a}=\underline{\alpha})$.

We shall consider in some detail the conditional properties of estimators of Types A and B. In addition, we give some results that may prove useful in the numerical evaluation of the conditional mean squared errors of other estimators of Types C and D. Estimators of Type E will not be considered.

As discussed in Chapter II, the balanced one-way random model can be interpreted as a Bayesian formulation of the corresponding balanced one-way fixed model, so that it is to be expected that the optimal unconditional estimators derived under the random model will be attractive under the fixed model from a Bayesian point of view. We shall show that they are also attractive from a classical or frequentist point of view. We find that, in general, estimators of $\mu_1,\ldots,\mu_I$ that have good unconditional properties when considered individually tend also to have good ensemble conditional properties.

We now describe some relationships between unconditional and conditional properties. Let $\underline{\hat{\mu}} = (\hat{\mu}_1,\ldots,\hat{\mu}_I)'$ denote an estimator of $\underline{\mu} = (\mu_1,\ldots\mu_I)'$ such that $E(\hat{\mu}_i)$ exists $(i = 1,\ldots,I)$. From standard properties of the expectation operator, we have that

$$E(\hat{\mu}_i-\mu_i) = E[CE(\hat{\mu}_i-\mu_i)], \qquad (V.A.1)$$

$$MSE(\hat{\mu}_i,\mu_i) = E[CMSE(\hat{\mu}_i,\mu_i)], \qquad (V.A.2)$$

$$TB(\underline{\hat{\mu}},\underline{\mu}) = E[TCB(\underline{\hat{\mu}},\mu)], \qquad (V.A.3)$$

and

$$TMSE(\underline{\hat{\mu}},\underline{\mu}) = E[TCMSE(\underline{\hat{\mu}},\underline{\mu})]. \qquad (V.A.4)$$

The following theorem describes some more substantial relationships.

Theorem V.A.1: Take the model to be the balanced one-way random model. Let $\underline{\hat{\mu}}_D = (\hat{\mu}_{D;1},\ldots,\hat{\mu}_{D;I})'$ be an arbitrary Type D estimator of $\underline{\mu}$. We have that

(i) $TCB(\underline{\hat{\mu}}_D,\underline{\mu}) = 0,$ (V.A.5)

and

(ii) $\mathrm{MSE}(\hat{\mu}_{D;i}, \mu_i) = (IJ)^{-1} E[\mathrm{TCMSE}(\underline{\hat{\mu}}_D, \underline{\mu})]$

$$(i = 1, \ldots, I). \qquad \text{(V.A.6)}$$

<u>Proof</u>: Let $\hat{\rho}_D$ denote the corresponding estimator of ρ. We have that

$$\begin{aligned}
\mathrm{TCB}(\underline{\hat{\mu}}_D, \underline{\mu}) &= J.\mathrm{CE}[\sum_i (\hat{\mu}_{D;i} - \mu_i)] \\
&= J.\mathrm{CE}\{\sum_i [\bar{y}_{i.} - \mu_i - \hat{\rho}_D(\bar{y}_{i.} - \bar{y}_{..})]\} \\
&= J\sum_i \mathrm{CE}(\bar{y}_{i.} - \mu_i) + J.\mathrm{CE}[\hat{\rho}_D \sum_i (\bar{y}_{i.} - \bar{y}_{..})] = 0,
\end{aligned}$$

which proves (i). Further, using Theorem IV.A.7 (i) and relationship (V.A.4), we find that

$$\begin{aligned}
\mathrm{MSE}(\hat{\mu}_{D;i}, \mu_i) &= (IJ)^{-1}\mathrm{TMSE}(\underline{\hat{\mu}}_D, \underline{\mu}) \\
&= (IJ)^{-1} E[\mathrm{TCMSE}(\underline{\hat{\mu}}_D, \underline{\mu})].
\end{aligned}$$

Q.E.D.

According to Theorem V.A.1, the components of a shrinkage estimator of $\underline{\mu}$ of Types A, B, C, or D are not necessarily conditionally unbiased, however, the sum of their conditional biases is zero. Further, conditional ensemble estimation is closely related to the unconditional estimation of individual components.

Let $\underline{\alpha}^* = (\alpha_1^*, \ldots, \alpha_I^*)'$, and recall that $s_\alpha^2 = (\sum_{i=1}^{I} \alpha_i^{*2})/(I-1) = [(\sum_{i=1}^{I} (\alpha_i - \bar{\alpha}_.)^2]/(I-1)$. Take M to be a random variable whose distribution is Poisson with parameter

$$\lambda = (1/2)[J(I-1)/\sigma_e^2]s_\alpha^2 = (1/2)(J/\sigma_e^2)\underline{\alpha}^{*\prime}\underline{\alpha}^*$$

unless $\lambda=0$, in which case define $M\equiv 0$. Note that $\lambda=0$ if and only if $\underline{\alpha}^*=\underline{0}$. Let w_1, w_2, w_3, w_4, and w_5 be random variables such that, conditional on M, $w_1 \sim \chi^2(IJ-I)$, $w_2 \sim \chi^2(I-1+2M)$, $w_3 \sim \chi^2(I+1+2M)$, $w_4 \sim \chi^2(I+1)$, and $w_5 \sim \chi^2(I-1)$, with w_1 being distributed independently of w_2, w_3, w_4, and w_5. Further, define s_2, s_3, s_4, and s_5 to be random variables such that, conditional on M, $s_2 \sim \text{beta}(\nu_e,\nu_2)$, $s_3 \sim \text{beta}(\nu_e,\nu_3)$, $s_4 \sim \text{beta}(\nu_e,\nu_4)$, and $s_5 \sim \text{beta}(\nu_e,\nu_5)$, with $\nu_e = (1/2)(IJ-I)$, $\nu_2 = (1/2)(I-1+2M)$, $\nu_3 = (1/2)(I+1+2M)$, $\nu_4 = (1/2)(I+1)$, and $\nu_5 = (1/2)(I-1)$.

We now give some results that are useful in computing conditional biases and mean squared errors for estimators of the first four types. These results are proven in the Appendix.

Let $\hat{\rho}_D = f_D(SS_e, SS_a)$, where f_D is an arbitrary positive function, and define $h_D(x,y) = f_D^2(x,y) - 2f_D(x,y)$ for $x>0$, $y>0$. Let $\hat{\underline{\mu}}_D$ and $\hat{\mu}_{D;i}$ be the corresponding Type D estimators of $\underline{\mu}$ and μ_i $(i = 1,\ldots,I)$. Take the model to be the balanced one-way random model.

If $CE(\hat{\mu}_{D;i})$ exists, then the conditional bias of

$\hat{\mu}_{D;i}$ is given by

$$CE(\hat{\mu}_{D;i}-\mu_i) = -CE[\hat{\beta}_D(\bar{y}_{i.}-\bar{y}_{..})]$$

$$= -(\sigma_e^2/J)\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[2Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)], \quad \text{if } \underline{\alpha}^* \neq \underline{0},$$

$$= 0, \quad \text{if } \underline{\alpha}^* = \underline{0}$$

$$(i = 1,\ldots,I), \tag{V.A.7}$$

as proven in Theorem VIII.B.2 in the Appendix.

The conditional mean squared error of $\hat{\mu}_{D;i}$ is

$$CMSE(\hat{\mu}_{D;i},\mu_i) = CE[(\hat{\mu}_{D;i}-\mu_i)^2]$$

$$= (\sigma_e^2/J)E\{1+[I-1)/I]h_D(\sigma_e^2 w_1,\sigma_e^2 w_3) + (2M)\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}$$

$$\cdot\ [h_D(\sigma_e^2 w_1,\sigma_e^2 w_3) + 2f_D(\sigma_e^2 w_1,\sigma_e^2 w_2)]\}, \quad \text{if } \underline{\alpha}^* \neq \underline{0},$$

$$= (\sigma_e^2/J)E\{1+[(I-1)/I]h_D(\sigma_e^2 w_1,\sigma_e^2 w_4)\}, \quad \text{if } \underline{\alpha}^* = \underline{0}$$

$$(i = 1,\ldots,I), \tag{V.A.8}$$

as proven in Theorem VIII.B.3 in the Appendix.

Theorem VIII.B.6 gives the following result, which is useful in computing conditional mean squared errors of linear combinations of the $\hat{\mu}_{D;i}$'s:

$$CE[(\hat{\mu}_{D;i}-\mu_i)(\hat{\mu}_{D;i'}-\mu_{i'})]$$

$$= \alpha_i^*\alpha_{i'}^* E\{h_D(\sigma_e^2 w_1, \sigma_e^2 w_3) + \frac{\sigma_e^2}{J}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1} 2Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)\},$$

$$\text{if } \underline{\alpha}^* \neq \underline{0},$$

$$= 0, \qquad \text{if } \underline{\alpha}^* = \underline{0}$$

$$(i > i' = 1,\ldots,I) \tag{V.A.9}$$

(if $CE[(\hat{\mu}_{D;i}-\mu_i)(\hat{\mu}_{D;i'}-\mu_{i'})]$ exists).

According to Theorem VIII.B.4 in the Appendix, the total conditional mean squared error of $\hat{\underline{\mu}}_D$ is

$$TCMSE(\hat{\underline{\mu}}_D, \underline{\mu}) = J\sum_i CE[(\hat{\mu}_{D;i}-\mu_i)^2]$$

$$= \sigma_e^2 \, E\{I + w_2 h_D(\sigma_e^2 w_1, \sigma_e^2 w_2) + 4Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)\}. \tag{V.A.10}$$

For estimators of Types A, B, and C, the above expressions can be simplified. Let $\hat{\rho}_C = f_C(SS_e/SS_a)$, where f_C is an arbitrary positive function, and define $h_C(x) = f_C^2(x) - 2f_C(x)$ for $x>0$. Let $\hat{\underline{\mu}}_C$ and $\hat{\mu}_{C;i}$ be the corresponding Type C estimator of $\underline{\mu}$ and μ_i $(i = 1,\ldots,I)$. Then,

$$CE[(\hat{\mu}_{C;i}-\mu_i)] = -(\sigma_e^2/J)\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[2Mf_C(\frac{s_2}{1-s_2})], \quad \text{if } \underline{\alpha}^*\neq\underline{0},$$

$$= 0, \qquad \text{if } \underline{\alpha}^*=\underline{0}$$

$$(i = 1,\ldots,I) \tag{V.A.11}$$

(if $CE(\hat{\mu}_{C;i})$ exists);

$$CMSE(\hat{\mu}_{C;i},\mu_i) = (\sigma_e^2/J)E\{1+ \frac{(I-1)}{I}h_C(\frac{s_3}{1-s_3})$$

$$+ 2M\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}$$

$$\cdot\ [h_C(\frac{s_3}{1-s_3})+2f_C(\frac{s_2}{1-s_2})]\}, \quad \text{if } \underline{\alpha}^*\neq\underline{0},$$

$$= (\sigma_e^2/J)E\{1+ \frac{(I-1)}{I}\ h_C(\frac{s_4}{1-s_4})\}, \quad \text{if } \underline{\alpha}^*=\underline{0}$$

$$(i = 1,\ldots,I); \tag{V.A.12}$$

$$CE[(\hat{\mu}_{C;i}-\mu_i)(\hat{\mu}_{C;i'}-\mu_{i'})] = \alpha_i^*\alpha_{i'}^*\ E\{\frac{\sigma_e^2}{J}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}2Mf_C(\frac{s_2}{1-s_2})$$

$$+ h_C(\frac{s_3}{1-s_3})\}, \qquad \text{if } \underline{\alpha}^*\neq\underline{0},$$

$$= 0, \qquad \text{if } \underline{\alpha}^*=\underline{0}$$

$$(i>i' = 1,\ldots,I) \tag{V.A.13}$$

(if $CE[(\hat{\mu}_{C;i}-\mu_i)(\hat{\mu}_{C;i'}-\mu_{i'})]$ exists);

$$\mathrm{TCMSE}(\underline{\hat{\mu}}_C,\underline{\mu}) = \sigma_e^2 E\{I+(IJ-1+2M)[(1-s_2)h_C(\frac{s_2}{1-s_2})]$$

$$+ 4M\ f_C(\frac{s_2}{1-s_2})\}. \qquad \text{(V.A.14)}$$

Results (V.A.11)-(V.A.14) correspond to Corollaries VIII.B.3, VIII.B.1, VIII.B.4, and VIII.B.2, respectively (Appendix).

Note that the total conditional mean squared errors for estimators of the first four types depend on the α_i's only through s_α^2. Let

$$s_a^2 = \sum_i (a_i - I^{-1}\sum_h a_h)^2/(I-1). \qquad \text{(V.A.15)}$$

Unconditionally,

$$s_a^2 \sim [\sigma_a^2/(I-1)]\chi^2(I-1) \qquad \text{(V.A.16)}$$

which means, using (V.A.6), that $\mathrm{MSE}(\hat{\mu}_{D;i},\mu_i)$ is a gamma transform of $\mathrm{TCMSE}(\underline{\hat{\mu}}_D,\underline{\mu})$. The family of distributions $[\sigma_a^2/(I-1)]\chi^2(I-1)$ is complete as a function of σ_a^2, which guarantees that different $\mathrm{TCMSE}(\underline{\hat{\mu}}_D,\underline{\mu})$'s transform into different $\mathrm{MSE}(\hat{\mu}_{D;i},\mu_i)$'s.

For estimators that are not of Types A or B, $\hat{\rho}_D$ is, possibly, a very complicated function of SS_e and SS_a, so that it is very difficult to obtain any further analytical simplification of expressions (V.A.7)-(V.A.14). In general, the evaluation of these expressions for a given estimator will require numerical integration. More tractable expressions for estimators of Types A and B will be given in sections V.B and V.C.

Recall that the best linear conditionally unbiased estimator (BLCUE) of μ_i is $\tilde{\mu}_i = \bar{y}_{i.}$ $(i = 1,\ldots,I)$, which corresponds to the ordinary least squares (OLS) estimator. If $I \geq 3$ then the vector $\tilde{\underline{\mu}} = \bar{\underline{y}} = (\bar{y}_{1.},\ldots,\bar{y}_{I.})'$ is inadmissible as an estimator of $\underline{\mu}$ with respect to the TCMSE criterion, a result discovered and first proved by Stein (1956). James and Stein (1961) produced an explicit estimator of $\underline{\mu}$ with TCMSE uniformly smaller than $\text{TCMSE}(\bar{\underline{y}},\underline{\mu}) = \sigma_e^2 I$ $(I \geq 3)$; we derive a version of the James-Stein estimator in section V.B. Theorem VIII.B.5 in the Appendix gives a class of estimators that dominate $\bar{\underline{y}}$. Unfortunately, estimators that reduce the total conditional mean squared error may have component-wise conditional mean squared errors that are unappealing. In using an estimator of $\underline{\mu}$ that has attractive ensemble conditional properties, the statistician must be aware of the possibility of grossly misestimating some individual components. In section V.B, we consider this issue, which was also considered by Efron and Morris (1972a) in a different context.

B. Type A Estimators

We use the notation of the preceding section. The following theorem gives the conditional biases for Type A estimators of μ_i.

Theorem V.B.1: Let k be an arbitrary positive constant and let $\hat{\mu}_{A,k;i}$ be the corresponding Type A estimator of μ_i. Then, under the balanced one-way random model,

$$\begin{aligned} CE(\hat{\mu}_{A,k;i}-\mu_i) \\ &= -2(\sigma_e^2/J)\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}kE[MI(J-1)(I-3+2M)^{-1}], \\ &\qquad\qquad \text{if } \underline{\alpha}^*\neq\underline{0}, \\ &= 0, \qquad\qquad \text{if } \underline{\alpha}^*=\underline{0} \end{aligned}$$

$$(I>3;\ i = 1,\ldots,I). \tag{V.B.1}$$

Proof: Taking $f_C(\cdot)$ to be the function

$$f_C(x) = kx \qquad (x>0), \tag{V.B.2}$$

we find that

$$\begin{aligned} E[2Mf_C(\frac{s_2}{1-s_2})] &= 2E[Mk\ s_2(1-s_2)^{-1}] \\ &= 2k\ E[M\ \nu_e(\nu_2-1)^{-1}] \\ &= 2k\ E[MI(J-1)(I-3+2M)^{-1}]. \end{aligned}$$

The theorem then follows from result (V.A.11).

Q.E.D.

The total conditional bias of $\underline{\hat{\mu}}_{A,k}$ is zero (Theorem V.A.1), however, some of the individual biases of the components of $\underline{\hat{\mu}}_{A,k}$ can be quite large. Let us consider the maximum value that can be attained by an individual conditional bias. For fixed $\sum_h \alpha_h^{*2} = c^2 > 0$, maximizing the absolute value of expression (V.B.1) is equivalent to maximizing α_i^{*2} subject to the restrictions

$$\sum_h \alpha_h^* = 0 \quad \text{and} \quad \sum_h \alpha_h^{*2} = c^2. \tag{V.B.3}$$

The Lagrangian for this maximization problem is

$$L_i = \alpha_i^{*2} - b_1(\sum_h \alpha_h^{*2} - c^2) - b_2(\sum_h \alpha_h^*), \tag{V.B.4}$$

where b_1 and b_2 are the Lagrange multipliers. We have that

$$\begin{aligned} \frac{\partial L_i}{\partial \alpha_n^*} &= 2\alpha_i^* - 2b_1\alpha_i^* - b_2, \quad \text{if} \quad n=i, \\ &= 2b_1\alpha_n^* - b_2, \quad \text{if} \quad n \neq i \end{aligned}$$

$$(n = 1, \ldots, I). \tag{V.B.5}$$

Consider the system of equations consisting of the constraints (V.B.3) together with the equations obtained by equating the partial derivatives (V.B.5) to zero. This system has two solutions for $\alpha_1^*, \ldots, \alpha_I^*$, namely

$$\alpha_n^* = \pm c[(I-1)/I]^{1/2}, \qquad \text{if } n=i,$$

$$= \mp c[I(I-1)]^{1/2}, \qquad \text{if } n\neq i$$

$$(n = 1,\ldots,I) \qquad \text{(V.B.6)}$$

It can be shown that, provided $I>3$, either of the solutions (V.B.6) maximizes α_i^{*2} and hence the absolute value of the bias of $\hat{\mu}_{A,k;i}$. As a function of $\underline{\alpha}^{*\prime}\underline{\alpha}^* = c^2>0$, the maximum absolute bias is

$$2(\sigma_e^2/J)|c|^{-1}[(I-1)/I]^{1/2}kE[MI(J-1)(I-3+2M)^{-1}]. \qquad \text{(V.B.7)}$$

It can be shown that the limit of this expression as $c^2\to 0$ is zero. Starting at $c^2=0$, the expression increases to a maximum and then decreases asymptotically to zero, the value of the bias of the ordinary least squares estimator of μ_i.

The following theorem gives total conditional mean squared errors for Type A estimators of $\underline{\mu}$.

<u>Theorem V.B.2</u>: Let k be an arbitrary positive constant and let $\underline{\hat{\mu}}_{A,k}$ be the corresponding Type A estimator of $\underline{\mu}$. Then,

$$\begin{aligned}\text{TCMSE}(\underline{\hat{\mu}}_{A,k},\underline{\mu}) &= \sigma_e^2 I-\sigma_e^4 kI(J-1)[2(I-3)-k(IJ-I+2)]CE(1/SS_a)\\ &= \sigma_e^2 I-\sigma_e^2 kI(J-1)[2(I-3)-k(IJ-I+2)]E[(I-3+2M)^{-1}]\end{aligned}$$

$$(I>3). \qquad \text{(V.B.8)}$$

In particular,

$$\mathrm{TCMSE}(\underline{\hat{\mu}}_{A,k};\underline{\mu}) = \sigma_e^2 I - \sigma_e^2 kI(J-1)[2(I-3)-k(IJ-I+2)](I-3)^{-1},$$

$$\text{if } \underline{\alpha}^*=\underline{0} \quad (I>3).$$

Proof: Note that, from property (VIII.A.4) in the Appendix, $CE[(SS_a)^{-1}] = \sigma_e^{-2}E[(I-3+2M)^{-1}]$. Let the function f_C be as given by (V.B.2) and define $h_C(\cdot)$ as

$$h_C(x) = f_C^2(x)-2f_C(x) = k^2x^2-2kx \quad (x>0). \tag{V.B.9}$$

We find that

$$E[(1-s_2)h_C(\frac{s_2}{1-s_2})|M] = E[k^2s_2^2(1-s_2)^{-1}-2ks_2|M]$$

$$= \frac{k^2(IJ-I)(IJ-I+2)}{(I-3+2M)(IJ-1+2M)} - \frac{2k(IJ-I)}{(IJ-1+2M)} \tag{V.B.10}$$

and that

$$E[f_C(\frac{s_2}{1-s_2})|M] = E[ks_2(1-s_2)^{-1}|M] = k(IJ-I)(I-3+2M)^{-1}. \tag{V.B.11}$$

Combining results (V.A.14), (V.B.10), and (V.B.11) we obtain expression (V.B.8).

Q.E.D.

Expression (V.B.8) is minimized uniformly with respect to k by

$$k^* = (I-3)(IJ-I+2)^{-1}. \tag{V.B.12}$$

As discussed in section IV.B, the value (V.B.12) also minimizes $MSE(\hat{\mu}_{A,k;i},\mu_i)$ uniformly. The estimator $\hat{\underline{\mu}}_{JS} = \hat{\underline{\mu}}_{A,k*}$ is of the type proposed by James and Stein (1961) and we subsequently refer to it as the James-Stein estimator. For the James-Stein estimator, expression (V.B.8) simplifies to

$$TCMSE(\hat{\underline{\mu}}_{JS},\underline{\mu}) = \sigma_e^2 I-\sigma_e^4 I(J-1)(I-3)^2(IJ-I+2)^{-1}CE(1/SS_a)$$

$$(I>3). \qquad (V.B.13)$$

The second term of expression (V.B.13) represents the decrease in TCMSE achieved by using the James-Stein estimator to estimate $\underline{\mu}$ rather than the ordinary least squares estimator $\tilde{\underline{\mu}} = \overline{\underline{y}}$.

The following theorem gives the individual conditional mean squared errors.

<u>Theorem V.B.3</u>: Let k be an arbitrary positive constant and let $\hat{\mu}_{A,k;i}$ be the corresponding Type A estimator of μ_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$\mathrm{CMSE}(\hat{\mu}_{A,i;k},\mu_i)$$

$$= (\sigma_e^2/J)E\{1 + \frac{kI(J-1)}{(I-1+2M)}\ \frac{k(I-1)(IJ-I+2)}{I(I-3+2M)}$$

$$+ (2M)\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^{*})^{-1}(\frac{k(IJ-I+2)+4}{(I-3+2M)})]\},\quad \text{if } \underline{\alpha}^{*}\neq\underline{0},$$

$$= (\sigma_e^2/J)\{1+k(J-1)[k(IJ-I+2)(I-3)^{-1}-2]\},\ \text{if } \underline{\alpha}^{*}=\underline{0}$$

$$(I>3;\quad i = 1,\ldots,I). \qquad \text{(V.B.14)}$$

Proof: Take h_C to be as defined by (V.B.9). We find that

$$E[h_C(\frac{s_3}{1-s_3})|M] = E[k^2s_3^2(1-s_3)^{-2}-2ks_3(1-s_3)^{-1}|M]$$

$$= kI(J-1)(I-1+2M)^{-1}[k(IJ-I+2)(I-3+2M)^{-1}-2]. \qquad \text{(V.B.15)}$$

Substituting expressions (V.B.15) and (V.B.11) into formula (V.A.11) for the case $\underline{\alpha}^{*}\neq\underline{0}$, we obtain, after some algebraic manipulation, formula (V.B.14). The derivation of formula (V.B.14) in the case $\underline{\alpha}^{*}=0$ proceeds along similar lines.

Q.E.D.

For the James-Stein estimator, the formula (V.B.14) becomes

$$\begin{aligned}
&\mathrm{CMSE}(\hat{\mu}_{JS;i},\mu_i) \\
&= (\sigma_e^2/J)E\{1 + \frac{(I-3)I(J-1)}{(IJ-I+2)(I-1+2M)}[\frac{(I-1)(I-3)}{I(I-3+2M)} \\
&\quad + (2M)\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)\frac{(I+1)}{(I-3+2M)}]\}, \qquad \text{if } \underline{\alpha}^*\neq\underline{0}, \\
&= (\sigma_e^2/J)\{1-(I-3)(J-1)/(IJ-I+2)\}, \qquad \text{if } \underline{\alpha}^*=\underline{0} \\
&\qquad\qquad (I>3;\quad i = 1,\ldots,I), \qquad\qquad (\mathrm{V.B.16})
\end{aligned}$$

where $\hat{\mu}_{JS;i} = \hat{\mu}_{A,k^*;i}$. Recall that $\mathrm{CMSE}(\bar{y}_{i.},\mu_i) = \sigma_e^2/J$. While the James-Stein estimator necessarily has smaller TCMSE than the ordinary least squares estimator $\bar{\underline{y}}$, the conditional mean squared error of the ith component of the James-Stein estimator may be much larger than that of $\bar{y}_{i.}$ for "unusually" large or small values of μ_i. For fixed $\underline{\alpha}^{*\prime}\underline{\alpha}^* = c^2 > 0$, the maximization of expression (V.B.16) [like the maximization of the more general expression (V.B.14) or the maximization of the absolute value of expression (V.B.1)] is equivalent to the maximization of α_i^{*2} subject to the restrictions (V.B.3). At the maximizing values of $\alpha_1^*,\ldots,\alpha_I^*$, the value of expression (V.B.16) becomes

$$(\sigma_e^2/J)E\{1 + \frac{(I-3)(J-1)(I-1)[I-3+2M(I+1)]}{(IJ-I+2)(I-1+2M)(I-3+2M)}\}. \qquad (\mathrm{V.B.17})$$

Starting at $\underline{\alpha}^{*\prime}\underline{\alpha}^* = 0$, the quantity (V.B.17), as a function of $\underline{\alpha}^{*\prime}\underline{\alpha}^*$, increases from its minimum value [given by expression (V.B.16), $\underline{\alpha}^*=\underline{0}$] to its maximum value and then decreases asymptotically to σ_e^2/J, the CMSE of the ordinary least squares estimator. The above discussion parallels that of Efron and Morris (1972a), although the context is different. These authors propose an estimator that can be viewed as a compromise between $\bar{\underline{y}}$ and the James-Stein estimator. Their compromise limits the conditional mean squared error of the individual components, while sacrificing a small fraction of the savings in total conditional mean squared error achieved by the James-Stein estimator.

The following theorem can be useful in finding conditional mean squared errors of Type A estimator of linear combinations of $\mu_1,\ldots\mu_I$.

<u>Theorem V.B.4</u>: Let k be an arbitrary positive constant and let $\hat{\mu}_{A,k;i}$ be the corresponding Type A estimator of μ_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$
\begin{aligned}
&CE[(\hat{\mu}_{A,k;i}-\mu_i)(\hat{\mu}_{A,k;i'}-\mu_{i'})] \\
&= \alpha_i^*\alpha_{i'}^* E\{\frac{kI(J-1)}{(I-1+2M)}[\frac{k(IJ-I+2)}{(I-3+2M)} - 2] \\
&\quad + \frac{2\sigma_e^2}{J}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}\frac{kMI(J-1)}{(I-3+2M)}\}, \qquad \text{if } \underline{\alpha}^*\neq\underline{0}, \\
&= 0, \qquad \text{if } \underline{\alpha}^*=\underline{0}
\end{aligned}
$$

$$(I>3; \quad i>i' = 1,\ldots,I). \qquad \text{(V.B.18)}$$

Theorem V.B.4 can be derived from formula (V.A.12) by making use of results (V.B.15) and (V.B.11).

C. Type B Estimators

We use the notation of section V.A. The following theorem gives an expression (in terms of incomplete beta function ratios) for the conditional bias of a Type B estimator of μ_i.

Theorem V.C.1: Let k be an arbitrary positive constant and let $\hat{\mu}_{B,k;i}$ be the corresponding Type B estimator of μ_i $(i = 1,\ldots,I)$. Define $z = z(k) = (1+k)^{-1}$. Then, under the balanced one-way random model,

$$\begin{aligned} CE(\hat{\mu}_{B,k;i}-\mu_i) &= -2(\sigma_e^2/J)\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E\{M[1-I_z(\nu_e,\nu_2) \\ &\qquad + \nu_e(\nu_2-1)^{-1}I_z(\nu_e+1,\nu_2-1)]\}, \qquad \text{if } \underline{\alpha}^*\neq\underline{0}, \\ &= 0, \qquad \text{if } \underline{\alpha}^*=\underline{0} \end{aligned}$$

$$(I>3;\quad i = 1,\ldots,I). \qquad \text{(V.C.1)}$$

Proof: Let

$$f_C(x) = \min\{kx, 1\} \qquad (x>0). \qquad \text{(V.C.2)}$$

The theorem follows from result (V.A.11) upon noting that

$$E[f_C(\frac{s_2}{1-s_2})|M] = E\{\min[ks_2(1-s_2)^{-1},1]|M\}$$

$$= \{\nu_e(\nu_2-1)^{-1}I_z(\nu_e+1,\nu_2-1)+1-I_z(\nu_e,\nu_2)\}. \qquad (V.C.3)$$

Q.E.D.

As in the case of Type A estimators, the total conditional bias of a Type B estimator of $\underline{\mu}$ is zero, but some of the individual biases can be very large, especially for unusually large or small values of μ_i.

<u>Theorem V.C.2</u>: Let k be an arbitrary positive constant and let $\hat{\mu}_{B,k}$ be the corresponding Type B estimator of $\underline{\mu}$. Then, under the balanced one-way random model,

$$\begin{aligned} \mathrm{TCMSE}(\underline{\hat{\mu}}_{B,k},\underline{\mu}) &= (1/2)\sigma_e^2(I+1)-(3/2)\sigma_e^2(I-1)(1-r)/r \\ &+ \sigma_e^2E\{I_z(\nu_e,\nu_2)[\frac{(\nu_e+1)^2\nu_e}{\nu_2(\nu_2-1)}k^2-2k\nu_e+\nu_2 \\ &+ 4kM\nu_e(\nu_2-1)^{-1}-4M] \\ &+ \frac{z^{\nu_e}(1-z)^{\nu_2}}{B(\nu_e,\nu_2)}[\frac{k^2(\nu_e+1)(2z\nu_e+z-z\nu_2-\nu_e-1)}{(1-z)\nu_2(\nu_2-1)} \\ &- 2k+1-\frac{4kM}{(1-z)(\nu_2-1)}]\}, \qquad (V.C.4)\end{aligned}$$

where $r = \sigma_e^2/(\sigma_e^2+Js_\alpha^2)$ and z is as defined in Theorem V.C.1.

<u>Proof</u>: Define

$$h_C(x) = f_C^2(x) - 2f_C(x) \quad (x>0), \qquad \text{(V.C.5)}$$

where f_C is given by (V.C.2). We find that

$$E\{(1-s_2)h_C(\frac{s_2}{1-s_2})|M\}$$

$$= E\{(1-s_2)\min[k^2 s_2^2(1-s_2)^{-2},1]$$

$$- 2(1-s_2)\min[ks_2(1-s_2)^{-1},1]|M\}$$

$$= k^2(\nu_e+1)\nu_e(\nu_2-1)^{-1}(\nu_e+\nu_2)^{-1}I_z(\nu_e+2,\nu_2-1)$$

$$- \nu_2(\nu_e+\nu_2)^{-1}[1-I_z(\nu_e,\nu_2+1)]$$

$$- 2\nu_e(\nu_e+\nu_2)^{-1}I_z(\nu_e+1,\nu_2). \qquad \text{(V.C.6)}$$

By substituting expressions (V.C.6) and (V.C.3) into formula (V.A.14), making use of some properties of the incomplete beta function ratio (e.g., Abramowitz and Stegun, 1970, relationships 26.5.11 and 26.5.16), and engaging in some algebraic manipulation, we obtain expression (V.C.4).
Q.E.D.

The following theorem establishes that a Type B (truncated) estimator of $\underline{\mu}$ always dominates the corresponding Type A (untruncated) estimator with respect to the TCMSE criterion.

Theorem V.C.3: Let $\hat{\rho}_U = f_U(SS_e/SS_a)$ where f_U is an arbitrary positive function and define $\hat{\rho}_T = \min\{\hat{\rho}_U, 1\}$. Take $\hat{\mu}_{U;i} = \bar{y}_{i.} - \hat{\rho}_U(\bar{y}_{i.} - \bar{y}_{..})$ and $\hat{\mu}_{T;i} = \bar{y}_{i.} - \hat{\rho}_T(\bar{y}_{i.} - \bar{y}_{..})$ $(i = 1,\ldots,I)$. Let $\underline{\hat{\mu}}_U = (\hat{\mu}_{U;1},\ldots,\hat{\mu}_{U;I})'$ and $\hat{\mu}_T = (\hat{\mu}_{T;1},\ldots,\hat{\mu}_{T;I})'$. Then, under the balanced one-way random model,

$$\mathrm{TCMSE}(\underline{\hat{\mu}}_T, \underline{\mu}) \leq \mathrm{TCMSE}(\underline{\hat{\mu}}_U, \underline{\mu}).$$

Proof: Let $f_T(x) \equiv \min\{f_U(x), 1\}$. Note that $f_T(x) = 1$, if $f_U(x) > 1$, and that $f_T(x) = f_U(x)$, otherwise. Using result (V.A.13), we have that

$$\begin{aligned}
&\mathrm{TCMSE}(\underline{\hat{\mu}}_U, \underline{\mu}) - \mathrm{TCMSE}(\underline{\hat{\mu}}_T, \underline{\mu}) \\
&\quad = \sigma_e^2 E\{(IJ-1+2M)(1-s_2)[f_U(\tfrac{s_2}{1-s_2})-1]^2 \\
&\qquad + 4M[f_U(\tfrac{s_2}{1-s_2})-1] \mid f_U(\tfrac{s_2}{1-s_2})>1\} \cdot P[f_U(\tfrac{s_2}{1-s_2})>1] \\
&\quad \geq 0.
\end{aligned}$$

Q.E.D.

As a particular case of the above theorem, we find that the truncated estimator $\underline{\hat{\mu}}_{TJS} = \underline{\hat{\mu}}_{B,k*}$ with k* given by (V.B.12) dominates the James-Stein estimator $(\underline{\hat{\mu}}_{JS})$ with respect to the TCMSE criterion. We refer to the estimator $\underline{\hat{\mu}}_{TJS}$ as the positive-part or truncated James-Stein estimator. The James-Stein estimator dominates $\underline{\bar{y}}$ and is the optimal Type A estimator of $\underline{\mu}$ with respect to the TCMSE criterion. Thus, since $\underset{\sim}{\hat{\mu}}_{TJS}$ dominates $\underset{\sim}{\hat{\mu}}_{JS}$, it dominates $\underset{\sim}{\bar{y}}$ and all Type A estimators.

The following theorem can be useful in deriving the conditional properties of Type B estimators of μ_i ($i = 1,\ldots,I$).

<u>Theorem V.C.4</u>: Let k be an arbitrary positive constant, and define $z = (1+k)^{-1}$. Let the function h_C be as defined by (V.C.5). Then, if $\nu_3>2$,

$$E\{h_C(\frac{s_3}{1-s_3})|M\} = \frac{k^2(\nu_e+1)\nu_e}{(\nu_3-2)(\nu_3-1)} I_z(\nu_e+2,\nu_3-2) - 1$$

$$+ I_z(\nu_e,\nu_3) - \frac{2k\nu_e}{(\nu_3-1)} I_z(\nu_e+1,\nu_e-1). \quad \text{(V.C.7)}$$

Theorem V.C.4 can be proved by using arguments similar to those used in the proof of Theorem V.C.2.

The quantities $CMSE(\hat{\mu}_{B,k;i},\mu_i)$ and $CE[(\hat{\mu}_{B,k;i}-\mu_i)(\hat{\mu}_{B,k;i'}-\mu_{i'})]$ $(i>i' = 1,\ldots,I)$ can be expressed in terms of incomplete beta function ratios by using expressions (V.A.12), (V.A.13), (V.C.3), and (V.C.7).

VI. REFERENCES

Abramowitz, M., and Stegun, I. A. 1970. Handbook of Mathematical Functions. National Bureau of Standards, Washington, D.C.

Aitken, A. C. 1934. "On Least Squares and Linear Combination of Observations." Proceedings of the Royal Society of Edinburgh 55:42-47.

Arnold, S. F. 1981. The Theory of Linear Models and Multivariate Analysis. Wiley, New York.

Baranchik, A. J. 1964. "Multiple Regression and Estimation of the Mean of a Multivariate Normal Distribution." Stanford University, Technical Report No. 51.

Baranchik, A. J. 1970. "A Family of Minimax Estimators of the Mean of a Multivariate Normal Distribution." The Annals of Mathematical Statistics 41:642-645.

Baranchik, A. J. 1973. "Inadmissibility of Maximum Likelihood Estimators in some Multiple Regression Problems with Three or More Independent Variables." The Annals of Statistics 1:312-321.

Bhattacharya, P. K. 1966. "Estimating the Mean of a Multivariate Normal Population with General Quadratic Loss Function." The Annals of Mathematical Statistics 37:1819-1824.

Box, G. E. P., and Tiao, G. C. 1973. Bayesian Inference in Statistical Analysis. Addison-Wesley, Reading, Massachusetts.

Brandwein, A., and Strawderman, W. 1980. "Minimax Estimators of Location Parameters for Spherically Symmetric Distributions with Concave Loss." The Annals of Statistics 8:64-74.

Chung, K. L. 1974. A Course in Probability Theory. Second Edition. Academic Press, New York.

Draper, N. R., and Van Nostrand, R. C. 1979. "Ridge Regression and James-Stein Estimation: Review and Comments." Technometrics 21:451-466.

Efron, B., and Morris, C. 1971. "Limiting the Risk of Bayes and Empirical Bayes Estimators - Part I: The Bayes Case." Journal of the American Statistical Association 66:807-815.

Efron, B., and Morris, C. 1972a. "Limiting the Risk of Bayes and Empirical Bayes Estimators - Part II: The Empirical Bayes Case." Journal of the American Statistical Association 67:130-139.

Efron, B., and Morris, C. 1972b. "Empirical Bayes on Vector Observations - An Extension of Stein's Method." Biometrika 59:335-347.

Efron, B., and Morris, C. 1973a. "Stein's Estimation Rule and its Competitors - An Empirical Bayes Approach." Journal of the American Statistical Association 68: 117-130.

Efron, B., and Morris, C. 1973b. "Combining Possibly Related Estimation Problems." Journal of the Royal Statistical Society, B 35:379-421.

Efron, B., and Morris, C. 1976a. "Families of Minimax Estimators of the Mean of a Multivariate Normal Distribution." The Annals of Statistics 4:11-21.

Efron, B., and Morris, C. 1976b. "Multivariate Empirical Bayes and Estimation of Covariance Matrices." The Annals of Statistics 4:22-32.

Ferguson, T. S. 1967. Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.

Graybill, F. A. 1976. Theory and Application of the Linear Model. Duxbury Press, North Scituate, Massachusetts.

Harville, D. A. 1969. "Variance-Component Estimation for the Unbalanced One-Way Random Classification - A Critique." Aerospace Research Laboratories, ARL69-0180.

Harville, D. A. 1976. "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects." The Annals of Statistics 4:384-395.

Harville, D. A. 1977. "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems." Journal of the American Statistical Association 72: 320-340.

Harville, D. A. 1978. "Alternative Formulations and Procedures for the Two-Way Mixed Model." Biometrics 34:441-453.

Harville, D. A. 1980. "Predictions for National Football League Games via Linear-Model Methodology." Journal of the American Statistical Association 75:516-524.

Hill, B. M. 1967. "Correlated Errors in the Random Model." Journal of the American Statistical Association 62: 1387-1400.

Hill, B. M. 1977. "Exact and Approximate Bayesian Solutions for Inference about Variance Components and Multivariate Inadmissibility." Pp. 129-152 in A. Aykac and C. Brumat, eds. New Developments in the Application of Bayesian Methods. North-Holland, Amsterdam.

Hill, B. M. 1980. "Robust Analysis of the Random Model and Weighted Least Squares Regression." Pp. 197-217 in J. Kmenta and J. Ramsey, eds. Evaluation of Econometric Models. Academic Press, New York.

James, W., and Stein, C. 1961. "Estimation with Quadratic Loss." Pp. 361-379 in J. Neyman, ed. Proceedings of the Fourth Berkeley Symposium, Vol. I. University of California Press, Berkeley.

Jeffreys, H. 1961. Theory of Probability. Third Edition, Claredon Press, Oxford.

Johnson, N. L., and Kotz, S. 1970. Continuous Univariate Distributions. Houghton-Mifflin, Boston.

Klotz, J. H., Milton, R. C., and Zacks, S. 1969. "Mean Square Efficiency of Estimators of Variance Components." Journal of the American Statistical Association 64: 1383-1402.

LaMotte, L. R. 1970. "A Class of Estimators of Variance Components." Technical Report No. 10. Department of Statistics, University of Kentucky.

LaMotte, L. R. 1971. "Locally Best Quadratic Estimators of Variance Components." Technical Report No. 22. Department of Statistics, University of Kentucky.

LaMotte, L. R. 1973. "Quadratic Estimation of Variance Components." Biometrics 29:311-330.

Naqvi, S. T. M. 1969. "Inference for Components of Variance Models." Unpublished Ph.D. Thesis. Library, Iowa State University.

Patel, J. K., Kapadia, C. H., and Owen, D. B. 1976. Handbook of Statistical Distributions. Marcek Dekker, Inc., New York.

Patterson, H. D., and Thompson, R. 1971. "Recovery of Inter-Block Information when Block Sizes are Unequal." Biometrika 58:545-554.

Patterson, H. D., and Thompson, R. 1974. "Maximum Likelihood Estimation of Components of Variance." Proceedings of the 8th International Biometric Conference, 197-207.

Rao, C. R. 1977. "Simultaneous Estimation of Parameters - A Compound Decision Problem." In S. S. Gupta and D. S. Moore, eds. Statistical Decision Theory and Related Topics II. Academic Press, New York.

Rao, C. R., and Shinozaki, N. 1979. "Precision of Individual Estimators in Simultaneous Estimation of Parameters." Biometrika 65:23-30.

Sclove, S. L. 1968. "Improved Estimators for Coefficients in Linear Regression." Journal of the American Statistical Association 63:596-606.

Sclove, S. L., Morris, C., and Radhakrishnan, R. 1972. "Nonoptimality of Preliminary-Test Estimators for the Multinormal Mean." The Annals of Mathematical Statistics 43:1481-1490.

Searle, S. R. 1971. Linear Models. John Wiley and Sons, New York.

Shinozaki, N. 1974. "A Note on Estimating the Mean Vector of a Multivariate Normal Distribution with General Quadratic Loss Function." Keio Engineering Reports 27:105-112.

Shinozaki, N. 1980. "Estimation of a Multivariate Normal Mean with a Class of Quadratic Loss Functions." Journal of the American Statistical Association 75: 973-976.

Snedecor, G. W., and Cochran, W. G. 1980. Statistical Methods. Seventh Edition. Iowa State University Press, Ames, Iowa.

Stein, C. 1956. "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution." Pp. 197-206 in J. Neyman, ed. Proceedings of the Third Berkeley Symposium, Vol. I. University of California Press, Berkeley.

Strawderman, W. E. 1971. "Proper Bayes Minimax Estimators of the Multivariate Normal Mean." The Annals of Mathematical Statistics 42:385-388.

Strawderman, W. E. 1973. "Proper Bayes Minimax Estimators of the Multivariate Normal Mean Vector for the Case of Common Unknown Variances." The Annals of Statistics 1:1189-1194.

Strawderman, W. E. 1978. "Minimax Adaptive Generalized Ridge Regression Estimators." Journal of the American Statistical Association 73:623-627.

Thisted, R. A. 1976. "Ridge Regression Minimax Estimation, and Empirical Bayes Methods." Technical Report No. 28. Division of Biostatistics, Stanford University.

Zacks, S. 1967. "More Efficient Estimators of Variance Components." Technical Report No. 4. Department of Statistics and Statistical Laboratory, Kansas State University.

VII. ACKNOWLEDGMENTS

I wish to thank my major professor, David A. Harville, for his continued encouragement and guidance throughout my period of study at Iowa State University.

I am extremely appreciative of the superb typing of this dissertation by Pat Gunnells.

Finally, I would like to express sincere appreciation to my wife for her encouragement and support over the years.

VIII. APPENDIX: EXPECTATIONS OF FUNCTIONS OF NONCENTRAL CHI-SQUARE RANDOM VARIABLES AND RELATED DEVELOPMENTS

A. General Results

The results presented in this section are an extension of results given by Shinozaki (1974), Baranchik (1973), and Arnold (1981, Chapter 11).

Let $\underline{u} = (u_1,\ldots,u_r)'$ be a normally distributed random vector with mean $\underline{\theta} = (\theta_1,\ldots,\theta_r)'$ and variance covariance matrix $\underline{I}$. Then, $\underline{u}'\underline{u} \sim \chi^2(r,\lambda)$, with $\lambda = (1/2)\underline{\theta}'\underline{\theta}$.

Define M to be a Poisson random variable with parameter $\lambda(\lambda\geq 0)$ (when $\lambda=0$ take $M\equiv 0$). Let v represent a random variable whose joint distribution, with M, is such that the conditional distribution of v given M is $\chi^2(r+2M)$. Then, the p.d.f. of the marginal distribution of v is

$$h(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \chi^2(x;r+2n) \qquad \text{(VIII.A.1)}$$

where $\chi^2(\cdot;\nu)$ denotes the p.d.f. of a central chi-square distribution with ν degrees of freedom.

Let $f(\cdot)$ be a measurable positive function such that

i) $E[|u_i|f(\underline{u}'\underline{u})]<\infty$,

and

ii) $E[|u_i u_j| f(\underline{u}'\underline{u})] < \infty$ $(i > j = 1, \ldots, r)$.

Let

$$g(M) = E[f(v)|M], \tag{VIII.A.2}$$

and let w represent a random variable whose distribution is $\chi^2(r+2)$. Then,

$$E[f(\underline{u}'\underline{u})] = E[g(M)] \tag{VIII.A.3}$$

In particular,

$$E[(\underline{u}'\underline{u})^{-1}] = E[(r+2M-2)^{-1}] \quad (r>2), \tag{VIII.A.4}$$

and

$$E[(\underline{u}'\underline{u})^{-2}] = E[(r+2M-2)^{-1}(r+2M-4)^{-1}] \quad (r>4). \tag{VIII.A.5}$$

An important recursive property of the Poisson distribution is

$$E[Mg(M)] = \lambda E[g(M+1)]. \tag{VIII.A.6}$$

Result (VIII.A.6) can be verified by the following argument:

$$E[Mg(M)] = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} ng(n)$$

$$= \sum_{n=1}^{\infty} \frac{e^{-\lambda}\lambda^n}{(n-1)!} g(n)$$

$$= \lambda \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} g(n+1)$$

$$= \lambda E[g(M+1)].$$

The following theorem gives a very useful result:

Theorem VIII.A.1: We have that

$$E[u_i f(\underline{u}'\underline{u})] = (\theta_i/\lambda)E[Mg(M)] = \theta_i E[g(M+1)], \quad \text{if } \lambda \neq 0,$$

$$= 0, \quad \text{if } \lambda = 0$$

$$(i = 1,\ldots,r). \tag{VIII.A.7}$$

Proof: Let $n(\underline{u}: \underline{\theta}, \underline{I})$ represent the density of $\underline{u}$. For $\lambda \neq 0$, we have that

$$\frac{\partial}{\partial\theta_i} E[f(\underline{u}'\underline{u})] = \frac{\partial}{\partial\theta_i} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\underline{u}'\underline{u})n(\underline{u}:\underline{\theta},\underline{I})d\underline{u}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\underline{u}'\underline{u})[\frac{\partial}{\partial\theta_i}n(\underline{u}:\underline{\theta},\underline{I})]d\underline{u}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\underline{u}'\underline{u})(u_i-\theta_i)n(\underline{u}:\underline{\theta},\underline{I})d\underline{u}$$

$$= E[(u_i-\theta_i)f(\underline{u}'\underline{u})] \tag{VIII.A.8}$$

and that

$$\frac{\partial}{\partial\theta_i} E[g(M)] = \frac{\partial}{\partial\theta_i} \sum_{n=0}^{\infty} g(n) \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= \sum_{n=0}^{\infty} g(n) [\frac{\partial}{\partial\theta_i} \frac{e^{-\lambda}\lambda^n}{n!}]$$

$$= \theta_i \sum_{n=0}^{\infty} g(n)(n/\lambda-1)\frac{e^{-\lambda}\lambda^n}{n!}$$

$$= (\theta_i/\lambda) \sum_{n=0}^{\infty} g(n)(n-\lambda)\frac{e^{-\lambda}\lambda^n}{n!}$$

$$= (\theta_i/\lambda)E[g(M)(M-\lambda)]. \qquad \text{(VIII.A.9)}$$

Then, combining results (VIII.A.3), (VIII.A.8), (VIII.A.9), and (VIII.A.6), we obtain formula (VIII.A.7) for the case $\lambda \neq 0$. If $\lambda=0$, then, given $\underline{u}'\underline{u}$, $u_i f(\underline{u}'\underline{u})$ is an odd integrable function of u_i; thus

$$E[u_i f(\underline{u}'\underline{u}) | \underline{u}'\underline{u}] = 0,$$

which implies

$$E[u_i f(\underline{u}'\underline{u})] = 0.$$

Q.E.D.

The following results can be proved by arguments analogous to those used to establish result (VIII.A.7);

$$E[u_i^2 f(\underline{u}'\underline{u})] = E[(1+\theta_i^2 M/\lambda)g(M+1)], \quad \text{if } \lambda \neq 0,$$

$$= E[f(w)], \qquad \text{if } \lambda=0$$

$$(i = 1,\dots,r); \qquad \text{(VIII.A.10)}$$

$$E[u_i u_j f(\underline{u}'\underline{u})] = \frac{\theta_i\theta_j}{\lambda} E[Mg(M+1)] = \theta_i\theta_j E[g(M+2)],$$

$$\text{if } \lambda \neq 0;$$

$$= 0, \qquad \text{if } \lambda = 0$$

$$(i>j = 1,\ldots,r). \qquad \text{(VIII.A.11)}$$

As particular cases of (VIII.A.10), we have that

$$E[u_i^2(\underline{u}'\underline{u})^{-1}] = E[(1+\theta_i^2 M/\lambda)(r+2M)^{-1}], \quad \text{if } \lambda \neq 0,$$

$$= 1/r, \qquad \text{if } \lambda = 0$$

$$(i = 1,\ldots,r), \qquad \text{(VIII.A:12)}$$

and

$$E[u_i^2(\underline{u}'\underline{u})^{-2}] = E\{[1+\theta_i^2 M/\lambda][r+2M)(r+2M-2)]^{-1}\},$$

$$\text{if } \lambda \neq 0,$$

$$= [r(r-2)]^{-1}, \qquad \text{if} \lambda = 0$$

$$(r>2;\ i = 1,\ldots,r). \qquad \text{(VIII.A.13)}$$

Two particular cases of (VIII.A.11) are:

$$E[u_i u_j(\underline{u}'\underline{u})^{-1}] = \frac{\theta_i\theta_j}{\lambda} E(\frac{M}{r+2M}) = \theta_i\theta_j E(\frac{1}{r+2M+2}),$$

$$\text{if } \lambda \neq 0,$$

$$= 0, \qquad \text{if } \lambda = 0$$

$$(i>j = 1,\ldots,r), \qquad \text{(VIII.A.14)}$$

and

$$E[u_i u_j (\underline{u}'\underline{u})^{-2}] = \frac{\theta_i\theta_j}{\lambda} E[\frac{M}{(r+2M)(r+2M-2)}]$$

$$= \theta_i\theta_j E[\frac{1}{(r+2M)(r+2M+2)}],$$

$$\text{if } \lambda \neq 0,$$

$$= 0, \qquad \text{if } \lambda = 0$$

$$(i>j = 1,\ldots,r). \qquad \text{(VIII.A.15)}$$

B. Applications to Conditional Properties under the Balanced One-Way Random Model

Take the model to be the balanced one-way random model. We use the notation introduced in section I.A.

Let $\underline{\bar{y}} = (\bar{y}_{1.},\ldots,\bar{y}_{I.})'$ and recall that $\underline{\mu} = (\mu_1,\ldots,\mu_I)'$. Then, conditional on $\underline{a}=\underline{\alpha}$, $\underline{\bar{y}} \sim N(\underline{\mu}, \frac{\sigma_e^2}{J}\underline{I})$. Define $\underline{1}$ to be an Ix1 vector of 1's, and take

$$\underline{P} = \underline{I} - (1/I)\underline{1}'\underline{1}. \qquad \text{(VIII.B.1)}$$

Note that $\underline{P}$ is a symmetric idempotent matrix of rank (I-1) and that

$$SS_a = J\sum_{i=1}^{I}(\bar{y}_{i.}-\bar{y}_{..})^2 = J\underline{\bar{y}}'\underline{P}\underline{\bar{y}}\ . \qquad \text{(VIII.B.2)}$$

Let $\underline{T}$ be an orthogonal IxI matrix such that

$$\underline{T}'\underline{PT} = \begin{bmatrix} \underline{I} & \underline{0} \\ \underline{0} & 0 \end{bmatrix}. \qquad \text{(VIII.B.3)}$$

Define

$$\underline{u} = (u_1,\ldots,u_I)' = (\sqrt{J}/\sigma_e)\underline{T}'\overline{\underline{y}}, \qquad \text{(VIII.B.4)}$$

and

$$\underline{\theta} = (\theta_1,\ldots,\theta_I)' = (\sqrt{J}/\sigma_e)\underline{T}'\underline{\mu}. \qquad \text{(VIII.B.5)}$$

Note that, conditional on $\underline{a}=\underline{\alpha}$, $\underline{u} \sim N(\underline{\theta},\underline{I})$. We also have that

$$\begin{aligned} SS_a &= J\overline{\underline{y}}'P\overline{\underline{y}} = \sigma_e^2\underline{u}'\underline{T}'\underline{PTu} \\ &= \sigma_e^2 \sum_{i=1}^{I-1} u_i^2. \end{aligned} \qquad \text{(VIII.B.6)}$$

Define

$$\underline{\Lambda} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1I} \\ \vdots & & \vdots \\ \lambda_{I1} & \cdots & \lambda_{II} \end{bmatrix} = (\sigma_e/\sqrt{J})\underline{PT} \qquad \text{(VIII.B.7)}$$

and note that $\lambda_{iI} = 0$ for $i = 1,\ldots,I$, i.e., the last column of $\underline{\Lambda}$ is null. We have that

$$\underline{P\overline{y}} = (\sigma_e/\sqrt{J})\underline{PTu} = \underline{\Lambda u} \qquad \text{(VIII.B.8)}$$

and hence,

$$\bar{y}_{i.}-\bar{y}_{..} = \text{ith component of } \underline{P}\bar{\underline{y}} = \sum_{j=1}^{I-1} \lambda_{ij}u_j$$

$$(i = 1,\ldots,I). \quad \text{(VIII.B.9)}$$

We have that

$$\underline{\Lambda\Lambda}' = (\sigma_e^2/J)\underline{PTT}'\underline{P} = (\sigma_e^2/J)\underline{P}, \quad \text{(VIII.B.10)}$$

implying that

$$\sum_{j=1}^{I-1} \lambda_{ij}^2 = \sum_{j=1}^{I} \lambda_{ij}^2 = (\sigma_e^2/J)(\text{ith diagonal element of P})$$
$$= \frac{\sigma_e^2(I-1)}{IJ} \qquad (i = 1,\ldots,I). \quad \text{(VIII.B.11)}$$

Also,

$$\underline{\Lambda\theta} = (\sigma_e/\sqrt{J})\underline{PT\theta} = \underline{P\mu}, \quad \text{(VIII.B.12)}$$

so that

$$\sum_{j=1}^{I=1} \lambda_{ij}\theta_j = \text{ith element of } \underline{P\mu} = \alpha_i-\bar{\alpha}_. = \alpha_i^*$$

$$(i = 1,\ldots,I). \quad \text{(VIII.B.13)}$$

Recall that $\underline{\alpha}^* = (\alpha_1^*,\ldots,\alpha_I^*)'$. Using result (VIII.B.3), we obtain

$$\sum_{h=1}^{I-1} \theta_h^2 = \underline{\theta}'\underline{T}'\underline{PT\theta} = (J/\sigma_e^2)\underline{\mu}'\underline{P\mu}$$

$$= (J/\sigma_e^2)\sum_{i=1}^{I} \alpha_i^{*2} = (J/\sigma_e^2)\underline{\alpha}^{*\prime}\underline{\alpha}^{*}. \qquad \text{(VIII.B.14)}$$

Result (VIII.B.14) implies that

$$\lambda = (1/2)(J/\sigma_e^2)\underline{\alpha}^{*\prime}\underline{\alpha}^{*} = (1/2)\sum_{h=1}^{I-1} \theta_h^2. \qquad \text{(VIII.B.15)}$$

We now use various results, together with the results of section A, to establish some useful theorems.

Let r = I-1 and let M, s_α^2, w_1, w_2, w_3, w_4, w_5, s_2, s_3, s_4, s_5, ν_e, ν_2, ν_3, ν_4, ν_5, and CE($\cdot$) be as defined in section V.A.

<u>Theorem VIII.B.1</u>: We have that, under the balanced one-way random model,

$$CE[(\bar{y}_{i.}-\bar{y}_{..})^2(SS_a)^{-2}]$$

$$= (\sigma_e^2 J)^{-1}\{\frac{(I-1)}{I} E[(I-1+2M)^{-1}(I-3+2M)^{-1}]$$

$$+ 2\alpha_i^{*2}[\underline{\alpha}^{*\prime}\underline{\alpha}^{*})^{-1}E[M(I-1+2M)^{-1}(I-3+2M)^{-1}]\},$$

$$\text{if } \underline{\alpha}^{*}\neq\underline{0},$$

$$= [\sigma_e^2 IJ(I-3)]^{-1}, \qquad \text{if } \underline{\alpha}^{*}=\underline{0}$$

$$(I>3;\quad i = 1,\ldots,I). \qquad \text{(VIII.B.16)}$$

Proof: Suppose that $\underline{\alpha}^*\neq\underline{0}$. Then,

$$CE[(\bar{y}_{i.}-\bar{y}_{..})^2(SS_a)^{-2}] = CE\{(\sum_{j=1}^{r}\lambda_{ij}u_j)^2(\sigma_e^2\sum_{h=1}^{r}u_h^2)^{-2}\}$$

$$= \sigma_e^{-4}\{\sum_{j=1}^{r}\lambda_{ij}^2 CE[u_j^2(\sum_{h=1}^{r}u_h^2)^{-2}]\}$$

$$+ 2\sum_{j=1}^{r}\sum_{\substack{\ell=1\\ \ell>j}}^{r}\lambda_{ij}\lambda_{i\ell}E[u_j u_\ell(\sum_{h=1}^{r}u_h^2)^{-2}]\}$$

$$= \sigma_e^{-4}\{E[(r+2M)^{-1}(r+2M-2)^{-1}](\sum_{j=1}^{r}\lambda_{ij}^2)$$

$$+ 2(\sum_{h=1}^{r}\theta_h^2)^{-1}(\sum_{j=1}^{r}\lambda_{ij}\theta_j)^2 E[M(r+2M)^{-1}(r+2M-2)^{-1}]\}$$

$$= (\sigma_e^2 J)^{-1}\{(I-1)I^{-1}E[(r+2M)^{-1}(r+2M-2)^{-1}]$$

$$+ 2\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[M(r+2M)^{-1}(r+2M-2)^{-1}]\},$$

which establishes result (VIII.B.16) for the case $\underline{\alpha}^*\neq\underline{0}$. A similar proof can be constructed for the case $\underline{\alpha}^*=\underline{0}$. Q.E.D.

Theorem VIII.B.2: Suppose that $\hat{\rho}_D = f_D(SS_e, SS_a)$ for some positive function f_D such that $CE[f_D(SS_e,SS_a)(\bar{y}_{i.}-\bar{y}_{..})]$ $(i = 1,\ldots,I)$ exists. Then, under the balanced one-way random model,

$$CE[\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})] = 2(\sigma_e^2/J)\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)],$$

$$\text{if } \underline{\alpha}^* \neq 0,$$

$$= 0, \qquad \text{if } \underline{\alpha}^* = \underline{0}$$

$$(i = 1,\ldots,I). \tag{VIII.B.17}$$

Proof: Suppose that $\underline{\alpha}^* = \underline{0}$. Let

$$g_1(c,M) = E[f_D(c,\sigma_e^2 v)\,|M], \tag{VIII.B.18}$$

where v is as defined in section A. Then,

$$CE[\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})] = E\{CE[f_D(SS_e,\sigma_e^2 \sum_{i=1}^{r} u_i^2)(\bar{y}_{i.}-\bar{y}_{..})\,|SS_e]\}$$

$$= \sum_{j=1}^{r} \lambda_{ij} E\{CE[f_D(SS_e,\sigma_e^2 \sum_{i=1}^{r} u_i^2)u_j\,|SS_e]\}$$

$$= (\sum_{j=1}^{r} \lambda_{ij}\theta_j)(\sum_{h=1}^{r} \theta_h^2)^{-1} E\{E[2Mg_1(SS_e,M)\,|SS_e]\}$$

$$= \alpha_i^*(\sigma_e^2/J)(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1} E\{E[2Mg_1(SS_e,M)\,|SS_e]\}$$

which is equivalent to result (VIII.B.17). The case $\underline{\alpha}^*=\underline{0}$ can be handled in a similar way.

Q.E.D.

Theorem VIII.B.3: Suppose that $\hat{\rho}_D = f_D(SS_e, SS_a)$ for some positive function f_D, and let $\hat{\mu}_{D;i}$ be the corresponding Type D estimator of μ_i. Then, under the balanced one-way random model,

$$CE[(\hat{\mu}_{D;i}-\mu_i)^2] = (\sigma_e^2/J)\{1+(I-1)/IE[h_D(\sigma_e^2 w_1,\sigma_e^2 w_3)]$$

$$+ 2\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[Mh_D(\sigma_e^2 w_1,\sigma_e^2 w_3)+2Mf_D(\sigma_e^2 w_1,\sigma_e^2 w_2)]\},$$

$$\text{if } \underline{\alpha}^*\neq 0,$$

$$= (\sigma_e^2/J)\{1+(I-1)/IE[h_D(\sigma_e^2 w_1,\sigma_e^2 w_4)],$$

$$\text{if } \underline{\alpha}^*=\underline{0}$$

$$(i = 1,\ldots,I), \qquad \text{(VIII.B.19)}$$

where

$$h_D(x,y) = f_D^2(x,y) - 2f_D(x,y). \qquad \text{(VIII.B.20)}$$

<u>Proof</u>: Suppose that $\underline{\alpha}^*\neq 0$. Then,

$$CE[(\hat{\mu}_{D;i}-\mu_i)^2] = CE[(\bar{y}_{i.}-\mu_i)^2]+CE[\hat{\rho}_D^2(\bar{y}_{i.}-\bar{y}_{..})^2]$$

$$- 2CE[\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i.}-\mu_i)]$$

$$= \sigma_e^2/J + CE[\hat{\rho}_D^2(\bar{y}_{i.}-\bar{y}_{..})^2] + 2CE[\alpha_i^*\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})]$$

$$= \sigma_e^2/J + CE[h_D(SS_e,SS_a)(\bar{y}_{i.}-\bar{y}_{..})^2]$$

$$+ 2\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}(\sigma_e^2/J)E[2Mf_D(\sigma_e^2 w_1,\sigma_e^2 w_2)] \qquad \text{(VIII.B,21)}$$

Further, using results (VIII.B.9), (VIII.A.7), (VIII.A.14), and (VIII.B.11), we find that

$$CE[h_D(SS_e,SS_a)(\bar{y}_{i.}-\bar{y}_{..})^2]$$

$$= E\{CE[h_D(SS_e,\sigma_e^2 \sum_{j=1}^{r} u_j^2)(\sum_{j=1}^{r} \lambda_{ij}u_j)^2 | SS_e]\}$$

$$= E\{E[\sum_{j=1}^{r} \lambda_{ij}^2 g_2(SS_e,M+1) + 2\sum_{\ell}\sum_{j>\ell} \lambda_{i\ell}(\sum_h \theta_h^2)^{-1}$$

$$\cdot Mg_2(SS_e,M+1) | SS_e]\}$$

$$= (\sigma_e^2/J)\{(I-1)I^{-1}E(E[g_2(SS_e,M+1)|SS_e])$$

$$+ 2\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^{*})^{-1}$$

$$\cdot E(E[Mg_2(SS_e,M+1)|SS_e])\}, \qquad \text{(VIII.B.22)}$$

where

$$g_2(c,\sigma_e^2 v) = E[h_D(c,\sigma_e^2 v)|M] \qquad \text{(VIII.B.23)}$$

with v as defined in section A. Combining results (VIII.B.21) and (VIII.B.22), we obtain result (VIII.B.19) for the case $\underline{\alpha}^*\neq\underline{0}$. A similar proof can be constructed for the case $\underline{\alpha}^*=\underline{0}$.

Q.E.D.

The following lemma is useful in our development:

Lemma VIII.B.1: Let $F(\cdot)$ represent an arbitrary measurable function, and let $w_6 \sim \chi^2(n)$ and $w_7 \sim \chi^2(n+2)$. Then,

$$E[nF(w_7)] = E[w_6F(w_6)]. \qquad \text{(VIII.B.24)}$$

Proof: We have that

$$E[nF(w_7)] = \int_0^\infty [2^{n/2+1}\Gamma(n/2+1)]^{-1} t^{n/2} e^{-t/2} nF(t)\,dt$$

$$= \int_0^\infty [2^{n/2}\Gamma(n/2)]^{-1} tF(t)\, t^{n/2-1} e^{-t/2} dt$$

$$= E[w_6 F(w_6)].$$

Q.E.D.

Theorem VIII.B.4: Suppose that $\hat{\rho}_D = f_D(SS_e, SS_a)$ for some positive function f_D and let $\hat{\mu}_D$ and $\hat{\mu}_{D;i}$ be the corresponding Type D estimators of $\underline{\mu}$ and μ_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$\begin{aligned} TCMSE(\underline{\hat{\mu}}_D, \underline{\mu}) &= J\sum_i CE(\hat{\mu}_{D;i}-\mu_i)^2 \\ &= \sigma_e^2\{I+E[h_D(\sigma_e^2 w_1, \sigma_e^2 w_2)w_2] \\ &\quad + 4E[Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)]\}, \end{aligned} \qquad \text{(VIII.B.25)}$$

where the function h_D is defined by (VIII.B.20).

This theorem can be proved by combining the results of Theorem VIII.B.3 and Lemma VIII.B.1).

Corollary VIII.B.1: Suppose that $\hat{\rho}_C = f_C(SS_e/SS_a)$ for some positive function f_C, and let $\hat{\mu}_{C;i}$ be the corresponding Type C estimator of μ_i $(i = 1,\ldots,I)$. Then, under the balanced one-way random model,

$$
\begin{aligned}
CE[(\hat{\mu}_{C;i}-\mu_i)^2]
&= (\sigma_e^2/J)E\{1+(I-1)I^{-1}h_C(\tfrac{s_3}{1-s_3}) + 2M\alpha_i^{*2}(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1} \\
&\quad \cdot [h_C(\tfrac{s_3}{1-s_3}) + 2f_C(\tfrac{s_2}{1-s_2})]\}, \quad \text{if } \underline{\alpha}^* \neq \underline{0}, \\
&= (\sigma_e^2/J)E\{1+(I-1)I^{-1}h_C(\tfrac{s_4}{1-s_4})\}, \quad \text{if } \underline{\alpha}^* = \underline{0}
\end{aligned}
$$

$$(i = 1,\ldots,I), \quad \text{(VIII.B.26)}$$

where

$$h_C(x) = f_C^2(x) - 2f_C(x). \quad \text{(VIII.B.27)}$$

Proof: Let $s_n = w_1/(w_1+w_n)$ $(n = 2,3,4)$. The distribution of s_n is beta with parameters ν_e and ν_n $(n = 2,3,4)$ (e.g., Johnson and Kotz, 1970, Section 24.2). Thus, the corollary follows from Theorem VIII.B.3.

Q.E.D.

Corollary VIII.B.2: Suppose that $\hat{\rho}_C = f_C(SS_e, SS_a)$ for some positive function f_C, and let $\underline{\hat{\mu}}_C$ be the corresponding Type C estimator of $\underline{\mu}$. Then, under the balanced one-way random model,

$$\mathrm{TCMSE}(\underline{\hat{\mu}}_C,\underline{\mu}) = \sigma_e^2 E\{I+(IJ-1+2M)[(1-s_2)h_C(\frac{s_2}{1-s_2})]$$

$$+ 4Mf_C(\frac{s_2}{1-s_2})\}, \qquad \text{(VIII.B.28)}$$

where the function h_C is defined by (VIII.B.27).

<u>Proof</u>: Let $s_2 = w_1/(w_1+w_2)$ and $z_2 = w_1+w_2$. Then, z_2 follows a central chi-square distribution with $I(J-1) + (I-1+2M) = IJ-1+2M$ degrees of freedom, s_2 follows a beta distribution with parameters ν_e and ν_2, and z_2 and s_2 are statistically independent (e.g., Johnson and Kotz, 1970, Section 24.2). We have that

$$E[h_C(w_1/w_2)w_2] = E\{h_C[s_2/(1-s_2)](1-s_2)z_2\}$$

$$= E(z_2)E\{h_C[s_2/(1-s_2)](1-s_2)\}$$

$$= (IJ-1+2M)E[h_C[s_2/(1-s_2)](1-s_2)\} \qquad \text{(VIII.B.29)}$$

and

$$E[f_C(w_1/w_2)] = E\{f_C[s_2/(1-s_2)]\}. \qquad \text{(VIII.B.30)}$$

The proof can now be completed by combining results (VIII.B.25), (VIII.B.29), and (VIII.B.30).

Q.E.D.

The following theorem is closely related to results given by Baranchik (1970) and Strawderman (1973).

Theorem VIII.B.5: Take the model to be the balanced one-way random model. Suppose that $\hat{\beta}_D = f_D(SS_e, SS_a)$, where f_D is an arbitrary positive function, and let $\underline{\hat{\mu}}_D$ be the corresponding Type D estimator of $\underline{\mu}$. Define $g_3(x,y) = yf_D(x,xy)$ for $x>0$ and $y>0$. If $g_3(\cdot,\cdot)$ satisfies the following conditions:

(i) for each fixed x, $g_3(x,\cdot)$ is monotone non-decreasing,

(ii) for each fixed y, $g_3(\cdot,y)$ is monotone non-increasing, and

(iii) $0 \leq g(\cdot,\cdot) \leq 2(I-3)/(IJ-I+2)$

then

$$TCMSE(\underline{\hat{\mu}}_D, \underline{\mu}) \leq \sigma_e^2 I.$$

Proof: Suppose that conditions (i), (ii), and (iii) are satisfied. We have that

$$f_D(\sigma_e^2 w_1, \sigma_e^2 w_2) = \frac{w_1}{w_2} g_3(\sigma_e^2 w_1, \frac{w_2}{w_1}) \leq \frac{w_1}{w_2} \frac{2(I-3)}{(IJ-I+2)} . \quad \text{(VIII.B.31)}$$

Thus,

$$E[f_D^2(\sigma_e^2 w_1, \sigma_e^2 w_2) - 2f_D(\sigma_e^2 w_1, \sigma_e^2 w_2) + 4Mf_D(\sigma_e^2 w_1, \sigma_e^2 w_2)]$$

$$= E\{g_3(\sigma_e^2 w_1, \frac{w_2}{w_1})[f_D(\sigma_e^2 w_1, \sigma_e^2 w_2)w_1 - 2w_1 + 4M \frac{w_1}{w_2}]\}$$

$$\leq E\{g_3(\sigma_e^2 w_1, \frac{w_2}{w_1})Z\}, \quad \text{(VIII.B.32)}$$

where

$$Z = w_1\{-2+[4M+2(I-3)(IJ-I+2)^{-1}w_1]/w_2\}.$$

According to Theorem VIII.B.4, it suffices to show that expression (VIII.B.32) does not exceed zero. Fixing w_1 and M, we define the constant a by

$$-2+[4M+2(I-3)(IJ-I+2)^{-1}w_1]/a = 0.$$

Note that $Z<0$ if $w_2>a$ and $Z\geq 0$ otherwise. From condition (i), we have the inequality

$$\begin{aligned}
&E\{g_3(\sigma_e^2 w_1, \tfrac{w_2}{w_1})Z|w_1, M\} \\
&\quad\leq g_3(\sigma_e^2 w_1, \tfrac{a}{w_1})E[Z|w_1,M;w_2\leq a]P(w_2\leq a|M) \\
&\quad+ g_3(\sigma_e^2 w_1, \tfrac{a}{w_1})E[Z|w_1,M;w_2>a]P(w_2>a|M) \\
&= g_3(\sigma_e^2 w_1, \tfrac{a}{w_1})E[Z|w_1,M] \\
&= g_3(\sigma_e^2 w_1, \tfrac{a}{w_1})w_1\{-2+[4M+2(I-3)(IJ-I+2)^{-1}w_1]/(I-3+2M)\} \\
&= g_3(\sigma_e^2 w_1, \tfrac{a}{w_1})w_1\{\frac{-(I-3+2M)+2M}{(I-3)} + \frac{w_1}{(IJ-I+2)}\}[\frac{2(I-3)}{(I-3+2M)}].
\end{aligned}$$

Therefore, it suffices to prove that

$$E\{g_3(\sigma_e^2 w_1, \frac{2M}{w_1} + \frac{I-3}{IJ-I+2})w_1[-1 + \frac{w_1}{IJ-I+2}]|M\} \qquad \text{(VIII.B.33)}$$

is less than or equal to zero. But, by conditions (i) and (ii), (VIII.B.33) is bounded above by

$$g_3[\sigma_e^2(IJ-I+2),\ \tfrac{2M+I-3}{IJ-I+2}]E\{w_1[-1+w_1/(IJ-I+2)]\,|w_1< IJ-I+2\}$$

$$\cdot\ P[w_1<IJ-I+2]+g_3[\sigma_e^2(IJ-I+2),\ \tfrac{2M+I-3}{IJ-I+2}]$$

$$\cdot\ E\{w_1[-1+w_1/(IJ-I+2)]\,|w_1 \geq IJ-I+2\}$$

$$\cdot\ P[w_1\geq IJ-I+2]$$

$$= g_3[\sigma_e^2(IJ-I+2),\ \tfrac{2M+I-3}{IJ-I+2}]E\{w_1[-1+w_1/(IJ-I+2]\} = 0$$

which completes the proof.

Q.E.D.

<u>Corollary VIII.B.3</u>: Suppose that $\hat{\rho}_C = f_C(SS_e/SS_a)$ for some positive function f_C such that $CE[\hat{\rho}_C(\bar{y}_{i.}-\bar{y}_{..})]$ $(i = 1,\ldots,I)$ exists. Then, under the balanced one-way random model,

$$\begin{aligned} CE[\hat{\rho}_C(\bar{y}_{i.}-\bar{y}_{..})] &= \sigma_e^2\alpha_i^*(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}E[2Mf_C(\tfrac{s_2}{1-s_2})], && \text{if } \underline{\alpha}^*\neq\underline{0}, \\ &= 0, && \text{if } \underline{\alpha}^*=\underline{0} \end{aligned}$$

$$(i = 1,\ldots,I). \qquad \text{(VIII.B.34)}$$

The proof of this corollary can be easily obtained from Theorem VIII.B.2, using arguments similar to those employed in the proof of Corollary VIII.B.1.

<u>Lemma VIII.B.2</u>: Take the model to be the balanced one-way random model. Suppose that $\hat{\rho}_D = f_D(SS_e, SS_a)$ for some positive function f_D. Then, if $CE\{\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\}$ exists,

$$CE\{\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\} = \alpha_i^*\alpha_{i'}^*E[f_D(\sigma_e^2 w_1, \sigma_e^2 w_3)]$$

$$(i>i' = 1,\ldots,I). \qquad \text{(VIII.B.35)}$$

<u>Proof</u>: Let $g_1(\cdot,\cdot)$ be as defined by (VIII.B.18). Using results (VIII.B.13) and (VIII.A.11) and proceeding as in the proof of Theorem VIII.B.2, we find that

$$CE\{\hat{\rho}_D(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\}$$

$$= CE\{\hat{\rho}_D \sum_{j=1}^{r}\sum_{j'=1}^{r}\lambda_{ij}\lambda_{i'j'}u_j u_{j'}\}$$

$$= \sum_j\sum_{j'}\lambda_{ij}\lambda_{i'j'}CE[\hat{\rho}_D u_j u_{j'}]$$

$$= \sum_j\sum_{j'}\lambda_{ij}\lambda_{i'j'}\theta_j\theta_{j'}E\{E[g_1(SS_e, M+2)\,|\,SS_e]\}$$

$$= \alpha_i^*\alpha_{i'}^*E\{E[g_1(SS_e, M+2)\,|\,SS_e]\}$$

which is equivalent to result (VIII.B.35).

Q.E.D.

Lemma VIII.B.3: Take the model to be the balanced one-way random model. Suppose that $\hat{\beta}_D = f_D(SS_e, SS_a)$ for some positive function f_D. Then, if $CE\{\hat{\beta}_D^2(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\}$ exists,

$$CE\{\hat{\beta}_D^2(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\} = \alpha_i^*\alpha_{i'}^* E[f_D^2(\sigma_e^2 w_1, \sigma_e^2 w_3)]$$

$$(i>i' = 1,\ldots,I). \qquad \text{(VIII.B.36)}$$

The proof of Lemma VIII.B.3 is very similar to that of Lemma VIII.B.2.

Lemma VIII.B.4: Take the model to be the balanced one-way random model. Suppose that $\hat{\beta}_D = f_D(SS_e, SS_a)$ for some positive function f_D. Then, if $CE\{\hat{\beta}_D(\bar{y}_{i.}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})\}$ exists,

$$\begin{aligned} CE\{\hat{\beta}_D(\bar{y}_{i.}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})\} \\ &= \alpha_i^*\alpha_{i'}^* E[f_D(\sigma_e^2 w_1, \sigma_e^2 w_3) - (\sigma_e^2/J)(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1} 2M f_D(\sigma_e^2 w_1, \sigma_e^2 w_2)], \\ &\qquad\qquad \text{if } \underline{\alpha}^* \neq \underline{0}, \\ &= 0, \qquad\qquad \text{if } \underline{\alpha}^* = \underline{0} \end{aligned}$$

$$(i>i' = 1,\ldots,I). \qquad \text{(VIII.B.37)}$$

<u>Proof</u>: Consider the case when $\underline{\alpha}^*\neq\underline{0}$. Then,

$$CE\{\hat{\beta}_D(\bar{y}_{i.}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})\}$$

$$= CE\{\hat{\beta}_D(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})\} + CE\{\hat{\beta}_D(\bar{y}_{..}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})\}$$

and

$$CE\{\hat{\beta}_D(\bar{y}_{..}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})\} = CE(\bar{y}_{..}-\mu_i)CE[\hat{\beta}_D(\bar{y}_{i'.}-\bar{y}_{..})]$$

$$= -\alpha_i^* CE[\hat{\beta}_D(\bar{y}_{i'.}-\bar{y}_{..})].$$

To complete the proof for the case when $\underline{\alpha}^*\neq\underline{0}$, we apply Theorem VIII.B.2 and Lemma VIII.B.2. The case when $\underline{\alpha}^*=\underline{0}$ can be handled in a similar way.

Q.E.D.

<u>Theorem VIII.B.6</u>: Take the model to be the balanced one-way random model. Suppose that $\hat{\beta}_D = f_D(SS_e,SS_a)$ for some positive function f_D and let $\hat{\mu}_{D;i}$ be the corresponding Type D estimator of μ_i $(i = 1,\ldots,I)$. Then, if $CE[(\hat{\mu}_{D;i}-\mu_i)(\hat{\mu}_{D;i'}-\mu_{i'})]$ exists,

$$\begin{aligned} CE[(\hat{\mu}_{D;i}-\mu_i)(\hat{\mu}_{D;i'}-\mu_{i'})] &= \alpha_i^*\alpha_{i'}^* E\{h_D(\sigma_e^2 w_1,\sigma_e^2 w_3) \\ &\quad + (\sigma_e^2/J)(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1} 2Mf_D(\sigma_e^2 w_1,\sigma_e^2 w_2)\}, \text{ if } \underline{\alpha}^*\neq\underline{0}, \\ &= 0, \qquad\qquad \text{if } \underline{\alpha}^*=\underline{0} \end{aligned}$$

$$(i>i' = 1,\ldots,I), \qquad \text{(VIII.B.38)}$$

where the function h_D is as defined by (A.B.20).

Proof: Consider the case when $\underline{\alpha}^*\neq\underline{0}$. Then,

$$CE[(\hat{\mu}_{D;i}-\mu_i)(\hat{\mu}_{D;i'}-\mu_{i'})]$$

$$= CE[(\bar{y}_{i.}-\mu_i)(\bar{y}_{i'.}-\mu_{i'})] + CE[\hat{\beta}_D^2(\bar{y}_{i.}-\bar{y}_{..})(\bar{y}_{i'.}-\bar{y}_{..})]$$

$$- CE[\hat{\beta}_D(\bar{y}_{i.}-\mu_i)(\bar{y}_{i'.}-\bar{y}_{..})]$$

$$- CE[\hat{\beta}_D(\bar{y}_{i'.}-\mu_{i'})(\bar{y}_{i.}-\bar{y}_{..})].$$

To complete the proof for the case when $\underline{\alpha}^*\neq\underline{0}$, apply Lemmas VIII.B.3 and VIII.B.4 and recall that $\bar{y}_{i..}$ and $\bar{y}_{i'..}$ are uncorrelated. The case when $\underline{\alpha}^*=\underline{0}$ can be handled in a similar way.

Q.E.D.

Corollary VIII.B.4: Take the model to be the balanced one-way random model. Suppose that $\hat{\beta}_C = f_C(SS_e/SS_a)$ for some positive function f_C, and let $\hat{\mu}_{C;i}$ be the corresponding Type C estimator of μ_i $(i = 1,\ldots,I)$. Then, if $CE[(\hat{\mu}_{C;i}-\mu_i)(\hat{\mu}_{C;i'}-\mu_{i'})]$ exists,

$$\begin{aligned} CE[(\hat{\mu}_{C;i}-\mu_i)(\hat{\mu}_{C;i'}-\mu_{i'})] &= \alpha_i^*\alpha_{i'}^* E[h_C(\tfrac{s_3}{1-s_3}) \\ &\quad + (\sigma_e^2/J)(\underline{\alpha}^{*\prime}\underline{\alpha}^*)^{-1}2Mf_C(\tfrac{s_2}{1-s_2})], &&\text{if } \underline{\alpha}^*\neq\underline{0}, \\ &= 0, &&\text{if } \underline{\alpha}^*=\underline{0} \end{aligned}$$

$$(i>i' = 1,\ldots,I), \qquad \text{(VIII.B.39)}$$

where the function h_C is as defined by (VIII.B.27).

Corollary VIII.B.4 follows almost immediately from Theorem VIII.B.6.